

Oracle ZFS Storage ZS3-BA

Installation and Configuration Guide



**ZFS STORAGE
APPLIANCE**

Part No.: E50598-02
June 2014

Copyright © 2012, 2014, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related software documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT END USERS. Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications which may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

Copyright © 2012, 2014, Oracle et/ou ses affiliés. Tous droits réservés.

Ce logiciel et la documentation qui l'accompagne sont protégés par les lois sur la propriété intellectuelle. Ils sont concédés sous licence et soumis à des restrictions d'utilisation et de divulgation. Sauf disposition de votre contrat de licence ou de la loi, vous ne pouvez pas copier, reproduire, traduire, diffuser, modifier, breveter, transmettre, distribuer, exposer, exécuter, publier ou afficher le logiciel, même partiellement, sous quelque forme et par quelque procédé que ce soit. Par ailleurs, il est interdit de procéder à toute ingénierie inverse du logiciel, de le désassembler ou de le décompiler, excepté à des fins d'interopérabilité avec des logiciels tiers ou tel que prescrit par la loi.

Les informations fournies dans ce document sont susceptibles de modification sans préavis. Par ailleurs, Oracle Corporation ne garantit pas qu'elles soient exemptes d'erreurs et vous invite, le cas échéant, à lui en faire part par écrit.

Si ce logiciel, ou la documentation qui l'accompagne, est concédé sous licence au Gouvernement des Etats-Unis, ou à toute entité qui délivre la licence de ce logiciel ou l'utilise pour le compte du Gouvernement des Etats-Unis, la notice suivante s'applique :

U.S. GOVERNMENT RIGHTS. Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle America, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

Ce logiciel ou matériel a été développé pour un usage général dans le cadre d'applications de gestion des informations. Ce logiciel ou matériel n'est pas conçu ni n'est destiné à être utilisé dans des applications à risque, notamment dans des applications pouvant causer des dommages corporels. Si vous utilisez ce logiciel ou matériel dans le cadre d'applications dangereuses, il est de votre responsabilité de prendre toutes les mesures de secours, de sauvegarde, de redondance et autres mesures nécessaires à son utilisation dans des conditions optimales de sécurité. Oracle Corporation et ses affiliés déclinent toute responsabilité quant aux dommages causés par l'utilisation de ce logiciel ou matériel pour ce type d'applications.

Oracle et Java sont des marques déposées d'Oracle Corporation et/ou de ses affiliés. Tout autre nom mentionné peut correspondre à des marques appartenant à d'autres propriétaires qu'Oracle.

Intel et Intel Xeon sont des marques ou des marques déposées d'Intel Corporation. Toutes les marques SPARC sont utilisées sous licence et sont des marques ou des marques déposées de SPARC International, Inc. AMD, Opteron, le logo AMD et le logo AMD Opteron sont des marques ou des marques déposées d'Advanced Micro Devices. UNIX est une marque déposée d'The Open Group.

Ce logiciel ou matériel et la documentation qui l'accompagne peuvent fournir des informations ou des liens donnant accès à des contenus, des produits et des services émanant de tiers. Oracle Corporation et ses affiliés déclinent toute responsabilité ou garantie expresse quant aux contenus, produits ou services émanant de tiers. En aucun cas, Oracle Corporation et ses affiliés ne sauraient être tenus pour responsables des pertes subies, des coûts occasionnés ou des dommages causés par l'accès à des contenus, produits ou services tiers, ou à leur utilisation.



Contents

Using This Documentation	vii
1. Introduction	1
Browser User Interface (BUI)	1
Command Line Interface (CLI)	2
2. Powering On and Initially Configuring the Appliance	3
Before You Begin	3
Prerequisites	3
Connecting to Oracle ILOM	4
▼ Logging In to Oracle ILOM Using a Local Serial Connection	5
▼ Logging In to Oracle ILOM Remotely Using a Web Interface	6
▼ Logging In to Oracle ILOM Remotely Using a Command Line Interface	6
Powering On the Appliance	7
Configuring the Primary Network Interface	7
Initially Configuring the Appliance	8
Initial Configuration	10
Post-Installation Controller Software Update	12
3. Connecting the Oracle ZFS Storage ZS3-BA	15

Oracle Exadata	15
Oracle SuperCluster T4-4	15
Connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T4-4	15
Connecting the Oracle ZFS Storage ZS3-BA Directly to the Oracle SuperCluster T4-4 InfiniBand Switches	16
Connecting the Oracle SuperCluster T4-4 External InfiniBand Leaf Switches	16
Oracle SuperCluster T5-8	17
Connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T5-8	17
Connecting the Oracle ZFS Storage ZS3-BA Directly to the Oracle SuperCluster T5-8 InfiniBand Switches	18
Connecting the Oracle ZFS Storage ZS3-BA to Oracle SuperCluster T5-8 External InfiniBand Leaf Switches	18
Oracle SuperCluster M6-32	19
Connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster M6-32	19

4. Configuring Oracle Exadata Backup 21

Automatic Configuration	21
Manual Configuration	22
General Configuration Guidelines	22
Configuring Networks, Pools, and Shares	22
Configuring Oracle RMAN and the Oracle Database Instance	24
Detailed Implementation Guidelines	25
Configuring the Oracle ZFS Storage ZS3-BA Network	25
Configuring the Oracle ZFS Storage ZS3-BA Storage Pool	31
Configuring the Oracle ZFS Storage ZS3-BA Cluster	32
Configuring the Oracle ZFS Storage ZS3-BA Share	33
Configuring the Oracle ZFS Storage ZS3-BA DTrace Analytics	35

Configuring the Client NFS Mount	36
Tuning the Linux Network and Kernel	37
Configuring Oracle Direct NFS (dNFS)	38
Tuning the Oracle Database Instance for Oracle RMAN Backup and Restore	39
Creating Dedicated Services for Oracle RMAN Operations	41
Configuring Oracle RMAN	42
5. Configuring Oracle SuperCluster Backup	49
Automatic Configuration	49
General Implementation Tasks	50
Detailed Implementation Tasks	50
Configuring the Oracle ZFS Storage ZS3-BA InfiniBand Datalinks	51
Configuring the Oracle SuperCluster InfiniBand Switches to Add the Oracle ZFS Storage ZS3-BA	52
Configuring Oracle ZFS Storage ZS3-BA Networking for Single IP Connection	55
Configuring Oracle ZFS Storage ZS3-BA Networking for an Active-Active Configuration	57
Configuring the Oracle ZFS Storage ZS3-BA Storage Pool	62
Configuring the Oracle ZFS Storage ZS3-BA Shares	62
Configuring the Oracle ZFS Storage ZS3-BA DTrace Analytics	65
Configuring the Client NFS Mount	66
Tuning the Solaris 11 Network and Kernel	66
Configuring Oracle Direct NFS (dNFS)	67
Tuning the Oracle Database Instance for Oracle RMAN Backup and Restore	68
Creating Dedicated Services for Oracle RMAN Operations	70
Configuring Oracle RMAN	71
6. Configuration Details for Client NFS Mount and Oracle dNFS	79

General Implementation Steps	79
Detailed Implementation Steps	80
Setting Up the Directory Structure to Mount the Shares on the Host	80
Updating a File to Mount Exported Shares	80
Updating the /etc/fstab File for Oracle Exadata	81
Updating the /etc/vfstab File for Oracle SuperCluster	82
Automating Share Mounting and Unmounting	83
Creating an init.d Service for Oracle Exadata	83
Enabling the NFS Client Service for Oracle SuperCluster	84
Updating orafstab to Access Oracle ZFS Storage ZS3-BA Exports	84
Updating orafstab to Access Oracle ZFS Storage ZS3-BA Exports for Oracle Exadata	85
Updating orafstab to Access Oracle ZFS Storage ZS3-BA Exports for Oracle SuperCluster	85
Mounting the Shares on the Host	86
Mounting the Shares on Oracle Exadata	86
Mounting the Shares on Oracle SuperCluster	87
Setting the Ownership of the Mounted Shares	87

Using This Documentation

This guide describes how to install, connect, and configure the Oracle ZFS Storage ZS3-BA for use with the Oracle Exadata Database Machine and the Oracle SuperCluster. It also contains procedures for setting up the Oracle Exadata and Oracle SuperCluster to use the Oracle ZFS Storage ZS3-BA as a backup target. For the latest version of this guide, visit

<http://www.oracle.com/goto/ZFSStorage/docs>.

Note – The Oracle ZFS Storage ZS3-BA is based on the Oracle ZFS Storage ZS3-4.

Product Notes

For late-breaking information and known issues about this product, visit My Oracle Support at <http://support.oracle.com>. Under the Knowledge tab, search for FAQ 1354980.1 for Oracle Exadata, or FAQ 1517107.1 for Oracle SuperCluster.

Related Documentation

Visit <http://www.oracle.com/goto/ZFSStorage/docs> for the following related documentation:

- *Oracle ZFS Storage Appliance Installation Guide*
- *Oracle ZFS Storage Appliance Customer Service Manual*
- *Oracle ZFS Storage Appliance Administration Guide*

■ *Oracle ZFS Storage Appliance Analytics Guide*

For white papers that describe using Oracle ZFS Storage ZS3-BA to back up Oracle Exadata and Oracle SuperCluster, visit

<http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/index.html>.

For a white paper that describes using Oracle ZFS Storage ZS3-BA to back up Oracle Database Appliance, visit

<http://www.oracle.com/technetwork/server-storage/engineered-systems/database-appliance/documentation/dbappliancebackupstrategies-519664.pdf>.

For information about using Oracle ILOM, refer to the Oracle Integrated Lights Out Manager (ILOM) 3.1 Documentation Library at

<http://www.oracle.com/pls/topic/lookup?ctx=ilom31>.

Feedback

Provide feedback about this documentation at:

<http://www.oracle.com/goto/docfeedback>

Access to Oracle Support

Oracle customers have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> or visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

Introduction

The Oracle ZFS Storage ZS3-BA, combined with Oracle 11gR2 Recovery Manager (Oracle RMAN), helps administrators meet the challenges of reliably backing up the Oracle Exadata Database Machine and Oracle SuperCluster 11gR2 in less time, thereby simplifying administration, and delivering smooth recovery operations when compromising conditions demand it. The cost-effective and high-bandwidth Oracle ZFS Storage ZS3-BA combines the simplicity of NFS protocol with ZFS-enhanced disk reliability. With native QDR InfiniBand (IB) and 10 gigabit (Gb) Ethernet connectivity, the Oracle ZFS Storage ZS3-BA is an ideal match for Oracle Exadata and Oracle SuperCluster. These high-bandwidth interconnects reduce backup and recovery time, as well as reduce backup application licensing and support fees, compared to traditional NAS storage systems.

Browser User Interface (BUI)

The BUI is the graphical tool for administration of the appliance. The BUI provides an intuitive environment for administration tasks, visualizing concepts, and analyzing performance data. The management software is designed to be fully featured and functional on a variety of web browsers.

Direct your browser to the appliance using either the IP address or host name you assigned to the NET-0 port during initial configuration as follows:
`https://ipaddress:215` or `https://hostname:215`. The login screen appears.

The online help link in the BUI is context-sensitive. For every top-level and second-level screen in the BUI, the associated help page appears when you click the Help button. The Oracle ZFS Storage ZS3-BA is based on the Oracle ZFS Storage ZS3-4.

Command Line Interface (CLI)

The CLI is designed to mirror the capabilities of the BUI, while also providing a powerful scripting environment for performing repetitive tasks. When navigating through the CLI, there are two principles to be aware of:

- Tab completion is used extensively: If you are not sure what to type in any given context, pressing the Tab key will provide you with possible options. Throughout the documentation, pressing Tab is presented as the word "tab" in bold italics.
- Help is always available: The help command provides context-specific help. Help on a particular topic is available by specifying the topic as an argument to help, for example help commands. Available topics are displayed by tab-completing the help command, or by typing help topics.

You can combine these two principles, as follows:

```
dory:> help tab
builtins      commands      general      help      properties      script
```

The Oracle ZFS Storage ZS3-BA is based on the Oracle ZFS Storage ZS3-4.

Powering On and Initially Configuring the Appliance

Before You Begin

This chapter includes instructions for powering on the appliance and configuring the primary network interface using the Oracle ILOM service processor (SP). Next, you configure the remaining system parameters and set up clustering using the browser user interface (BUI) from any client on the same network. Finally, perform the post-installation software update procedure to ensure you are running the latest software release.

Prerequisites

The appliance is configured and administered either directly using an administrative client (terminal or terminal emulator) connected to the serial management port on the controller, or indirectly via your network and the network management port. To access the appliance indirectly, your network must have a Dynamic Host Configuration Protocol (DHCP) server. Network access is preferred when connecting to the appliance for Oracle ILOM, discussed in the next section.

To prepare for connecting the appliance to your network, gather the following information:

- Host name
- Default router IP address
- IP address/netmask for management interface on primary controller
- IP address/netmask for management interface on peer controller

- IP address/netmask for data interface on primary controller (InfiniBand IPMP interface #1)
- IP address/netmask for data interface on peer controller (InfiniBand IPMP interface #2)
- DNS IP address #1/#2
- DNS IP address #3/#4
- DNS/NIS/LDAP domain name
- NTP server IP address
- LDAP: Use SSL/TLS? (Yes/No)
- LDAP server IP address #1/#2 (port)
- LDAP server IP address #3/#4 (port)
- LDAP base search DN
- LDAP search scope recursive? (Yes/No)
- LDAP authentication method: Simple, SASL/Digest MD5, or SASL/GSSAPI
- LDAP proxy DN (or anonymous)
- LDAP proxy DN password
- NIS server IP address #1/#2
- Desired administrative password for the appliance

To prepare for directly controlling the appliance via the serial port, configure an administrative client with the following settings:

- 8N1: eight data bits, no parity, one stop bit
- 9600 baud
- Disable hardware flow control (CTS/RTS)
- Disable software flow control (XON/XOFF)

Connecting to Oracle ILOM

In rare cases, faults associated with uncorrectable CPU errors are not diagnosable or displayed in the controller. These faults are preserved by and observable on the Oracle Integrated Lights Out Manager (ILOM), which resides in the Service Processor (SP). Connect to the server ILOM to diagnose hardware faults that do not appear in the BUI.

Since this appliance has two clustered controllers, an Oracle ILOM connection should be made to each controller. Each controller should be assigned a unique management interface so that it is not used as a cluster data interface. For example, the primary controller could use interface igb0 and the peer controller could use interface igb1 as the management interfaces. To enable interface management after

the appliance is configured, in the BUI, navigate to **Configuration > Network**, and edit the appropriate interface by clicking on its pencil icon. Then select the box **Allow Administration** and click **Apply**.

After the appliance is initially configured, you should lock the management interfaces, which makes them private network interfaces, to enable administration via either controller over the network (BUI or CLI) regardless of the cluster state. If interfaces are not locked and cluster takeover occurs, there is no way to access diagnostic information on the failed controller because interfaces are taken over by the other controller. A locked interface displays a lock icon next to its name in the BUI Configuration > Cluster screen. To lock interfaces, in the BUI, navigate to **Configuration > Cluster**. In the Resource list, choose the management interface for the controller and set it to a private, locked interface by clicking the unlock icon. The interface then displays a lock icon next to its name in the Resource list.

The server ILOM provides options for (i) network and (ii) serial port connectivity. Network connection is the preferred choice because the ILOM serial port does not always allow adequate means of platform data collection. Network connections are on port 215, and serial connections are on SSH port 22.



Caution – Failure to configure Oracle ILOM connectivity may lead to longer than necessary hardware fault diagnosis and resolution times.

For information about using Oracle ILOM, refer to the Oracle Integrated Lights Out Manager (ILOM) 3.1 Documentation Library at <http://www.oracle.com/pls/topic/lookup?ctx=ilom31>.

Choose the appropriate login procedure for your environment:

- Local Serial Connection
- Remote Connection Using a Web Interface
- Remote Connection Using a CLI

▼ Logging In to Oracle ILOM Using a Local Serial Connection

1. Prepare an administrative client (terminal or terminal emulator) with the serial connection parameters defined in the Prerequisites section.
2. Connect the administrative client to the SER MGT port on the controller using a serial null modem cable.
3. To establish a connection between your serial console and Oracle ILOM, press **Enter** on the administrative client.

A login prompt for Oracle ILOM is displayed.

4. Log in to the Oracle CLI using the administrative account name and password (defaults are **root** and **changeme**).

Oracle ILOM displays a default command prompt (->), indicating that you have successfully logged in.

▼ Logging In to Oracle ILOM Remotely Using a Web Interface

This procedure requires that you know the administrative account name and password (defaults are **root** and **changeme**), and that you know the IP address or host name of the controller Service Processor (SP). To improve response times, disable the web browser proxy server, if used.

1. Type the controller SP IP address into your web browser address field.

Example: **https://172.16.82.26**

2. Type the administrative user name and password.

3. Click Log In.

The System Information Summary page is displayed.

▼ Logging In to Oracle ILOM Remotely Using a Command Line Interface

This procedure requires that you know the administrative account name and password (defaults are **root** and **changeme**), and that you know the IP address or host name of the controller Service Processor (SP).

1. Using a Secure Shell (SSH) session, log in to Oracle ILOM by specifying your administrative account user name and the IP address or host name of the controller SP.

Examples:

```
ssh -l username host
```

```
ssh username@host
```

where *host* is either the IP address or host name of the controller SP when using DNS. For example: **ssh root@172.16.82.26**

2. Type the password for the administrative account.

Oracle ILOM displays a default command prompt (->), indicating that you have successfully logged in.

Powering On the Appliance

You will power on both controllers, but initially connect to and configure one controller.

1. **Connect to one controller (either one). This is referred to as the primary controller.**
 - To perform the initial configuration using a serial connection, connect a serial cable from the SER MGT port on the back panel of one controller to the serial port on the administrative client. Use a DB9 to RJ45 adapter if necessary.
 - To perform the initial configuration using an Ethernet connection, connect an Ethernet cable from the NET MGT port on the back panel of one controller to your network.
2. **Connect an Ethernet cable from your network to the NET-0 port on the back panel of the controller. Repeat for the second controller.**
3. **Power on the disk shelves attached to the storage system by plugging the power cords into the universal power connectors, connecting the cords to the external power source, and turning on the disk shelf power switches. Wait several minutes until the power indicators are lit a steady green.**
4. **Connect power cables to both power supplies on the storage controller and wait until the Power/OK LED on the front panel next to the Power button lights and remains lit (approximately two minutes). Repeat for the second controller.**

Configuring the Primary Network Interface

On the primary controller, connect to the Service Processor (SP) and configure the primary network interface with the network parameters you gathered in the Prerequisites section.

1. **Open a terminal window or terminal emulator and issue the appropriate command, as follows:**
 - For a serial port concentrator (for example, using telnet), connect and use **root** as the user name, and replace *serial-concentrator portnumber* with the appropriate value, for example: **telnet serial-concentrator portnumber**

- For a network connection, connect using SSH and use **root** as the user name, and determine the IP address by accessing your DHCP server for the address assigned to the Service Processor MAC address (see the label on the storage controller), for example: **ssh root@192.168.128.256**
 - a. Type the password **changeme** when prompted.
 - b. At the command prompt, enter **start /SP/console**.
 - c. Type **y** to confirm that you want to start the console.
- 2. Press any key to begin configuring the primary network interface. The shell interface configuration screen appears. The "=" character is displayed next to NET-0 at the top of the screen. If "X" appears by NET-0, check that all cable connections are tight.

To exit the shell interface before completing configuration, enter **exit**.
- 3. Verify the information on the screen, or enter values that do not appear. Apply the values by pressing ESC-1 or the F1 key.
- 4. After setting a new administrative password, press Enter to confirm the password and complete this stage of the configuration.

The final shell configuration screen appears, confirming that your appliance is ready for further configuration using the BUI.

Initially Configuring the Appliance

Configure the remaining system parameters and set up clustering on the primary controller through the BUI using a browser running on any client on the same network as the initial interface.

During cluster initialization, the software propagates the configuration from the primary controller to the peer controller. After the cluster is initialized, you can administer the system from either storage controller. However, do not attempt initial configuration on both controllers independently. Refer to the *Oracle ZFS Storage Appliance Administration Guide*

(<http://www.oracle.com/goto/ZFSStorage/docs>) for more information on setting up a clustered environment.

The management software is designed to be fully featured and functional on a variety of web browsers.

Note – Do not use Internet Explorer version 6 or earlier.

There are six steps for initially configuring the:

- Network
- DNS
- Time
- Name Services (NIS, LDAP, Active Directory)
- Storage
- Registration and Support

When configuring these parameters, follow these guidelines:

- Network interfaces configured via DHCP cannot be failed over between controllers and, therefore, cannot be used by clients to access storage. Be sure to assign static IP addresses to any network interfaces which will be used by clients to access storage, including, if applicable, the network interface used for configuration in section [“Powering On the Appliance” on page 7](#).

Note – This must be done before performing the initial configuration in this section.

- As a best practice, configure and assign a private network interface for administration to each controller, which enables administration via either controller over the network (BUI or CLI) regardless of the cluster state.
- **IMPORTANT:** If routes are needed, be sure to create a route on each interface that is assigned to a controller.
- Each storage pool can be taken over, along with the network interfaces clients use to reach that storage pool, by the cluster peer when takeover occurs.
- If you create two storage pools, each controller normally provides clients with access to the pool assigned to it. If one of the controllers fails, the other provides clients with access to both pools.
- If you create a single pool, the controller which is not assigned a pool provides service to clients only when its peer has failed.
- Storage pools are assigned to controllers at the time of creation; the storage configuration dialog offers the option of creating a pool assigned to each controller independently.
- The smallest unit of storage that may be assigned to a pool is one drive; however, it is recommended to use eight drives as a minimum and ideally more. Also, fewer pools with more disks per pool are preferred because they simplify management and provide a higher percentage of overall usable capacity.
- If you create two pools, there is no requirement that they must be the same size. Any subdivision of available storage is permitted.
- After completing basic configuration, you can assign resources to each controller. Typically, you assign only network interfaces because storage pools are automatically assigned during the storage configuration step.

For more information, especially clustering considerations, see the *Oracle ZFS Storage Appliance Administration Guide* (<http://www.oracle.com/goto/ZFSStorage/docs>).

Note – Be sure to follow the procedure at the end of this chapter, “[Post-Installation Controller Software Update](#)” on page 2-12, to ensure you are running the latest software release.

Initial Configuration

1. Direct your browser to the system using either the IP address or host name you assigned to the NET-0 port as follows: **`https://ipaddress:215`** or **`https://hostname:215`**. The login screen appears.

If the login screen does not appear and the message `Secure Connection Failed` is displayed, click the link to add an exception, download the certificate, and click `Confirm Security Exception`.

Note – Only perform this initial configuration on the primary controller.

2. Type **root** into the Username field, your administrative password, and press the Enter key. The Welcome screen appears.
3. To begin configuring the system, click **Start** on the Welcome screen.
4. Click **Cluster**. You are guided through the initial configuration, one screen at a time.
5. Click **Commit** to save the configuration and go to the next screen. Arrows beneath the Commit button can be used to revisit previous steps, and change the configuration if desired.
 - a. When a diagram of the active cluster links is displayed, ensure that three solid blue lines are shown. If not, ensure that the three cables connecting the two controllers together are tight in their connectors.
 - b. When prompted for the second controller’s information, enter its unique name and the default administrative password (**changeme**).

- c. When the Network configuration screen is displayed, add six interfaces: one management interface for each controller (typically igb0 and igb1) and four InfiniBand interfaces (typically ibp0, ibp1, ibp2, and ibp3). Create two IPMP groups for the InfiniBand interfaces, and check routing.

To add an interface, drag a device from the Devices column to the Datalinks column. Then drag it from the Datalinks column to the Interfaces column and add the IP address, along with other configuration data. As a best practice, lock the management interfaces to make them private network interfaces.

- i. When configuring the InfiniBand interfaces, use this IP address and mask: **0.0.0.0/8**. Select the checkbox **IB Partition**, enter the partition key **ffff**, and select **Connected Mode** for the link mode.
 - ii. To create the first IPMP group, click the “+” character next to Interfaces, select the checkboxes for **Enable Interface** and **Allow Administration**, add the IP address and netmask, and select the checkbox **IP MultiPathing Group**. Select the checkbox for ibp0 and select **Active** from the pull-down list. Select the checkbox for ibp3 and select **Passive** from the pull-down list.
 - iii. To create the second IPMP group, click the “+” character next to Interfaces, select the checkboxes for **Enable Interface** and **Allow Administration**, add the IP address and netmask, and select the checkbox **IP MultiPathing Group**. Select the checkbox for ibp1 and select **Passive** from the pull-down list. Select the checkbox for ibp2 and select **Active** from the pull-down list.
 - iv. Click the Routing submenu and ensure that the default route is using the first interface, which is typically igb0. Configure a second default gateway, which is typically igb1. All data interface routes must be configured manually; there are no default routes.
- d. When the DNS configuration screen is displayed, verify the settings against the parameters you gathered in the Prerequisites section.
- You must enter values for the DNS server. To enter additional DNS servers, click the “+” character and add them.
- e. When the NTP configuration screen is displayed, ensure that the correct server and client times are shown under the Clock heading, and enter the NTP server IP address.
- To adjust the time, enter the correct time for the client and click Sync.
- f. When the Directory Services configuration screen is displayed, click the pencil icon next to the desired service to edit its settings. Enter the parameters you gathered in the Prerequisites section.
- v. For NIS, instead of entering the IP addresses for the two NIS servers, you may select Use Broadcast if this matches your network.

- vi. For LDAP, select the checkbox to protect LDAP traffic with SSL/TLS, use the distinguished name (DN) in LDAP notation, select Anonymous or Proxy for the bind credential level and enter the DN and password of the account to be used if selecting Proxy, and add the certificate if configured and present. To add another LDAP server, click the “+” character next to Servers.
 - vii. For Active Directory, if selecting Join Workgroup, enter the workgroup name and click Apply. If selecting Join Domain and you have an AD site name or a preferred domain controller, add it to the CIFS configuration page. The user account to join the server to the domain must have sufficient privileges to do so. If the AD domain and the DNS domain do not match, enter the DNS domain in the field Additional DNS search path. Also ensure that the times on the appliance and the AD server match within five minutes.
 - viii. When the summary of Directory Services settings is displayed, click Commit if all settings are correct. Otherwise, click Return to Services and correct the settings.
 - g. When configuring storage, see the appropriate section in this guide for Oracle Exadata or Oracle SuperCluster.
6. After completing the Registration and Support step, the configuration is then propagated to the peer controller, including the new administrative password.

The initial configuration can be repeated at a later time by clicking the “INITIAL SETUP” button on the Maintenance > System screen.

Post-Installation Controller Software Update

Note – Use the following procedure to update your controller software to ensure you are running the latest software release.

Maintaining controller software in advance of putting your appliance into production enables you to reap the greatest benefits of recent software enhancements. In many cases, a simple software update will resolve an issue observed in testing or provide you with new enhancements that improve productivity. Update storage controllers to the latest software and associated firmware using the following procedure.

1. To view the controller software version, go to Maintenance > System. Continue with this procedure if an update is needed.
2. Ensure that any resilvering operations have completed by checking the Configuration > Storage screen.
3. Ensure that there are no active problems on the Maintenance > Problems screen.
4. Verify that hardware firmware updates are not in progress on the Maintenance > System screen.
5. Click the “Sign In” link at <http://support.oracle.com>.
The Oracle Single Sign-On screen appears.
6. Log in using your Oracle account credentials.
7. Click the Patches & Updates tab.
8. In the Search screen, select Product or Family (Advanced), which is near the top.
9. In the Product field, type and accept Sun ZFS Storage Appliance.
10. In the Release field, select the appropriate release from the drop-down list.
11. Click Search.
12. Select the desired release by clicking on the Patch Name.
The patch information, along with download links, is displayed.
13. View the Read Me file, which contains the Release Notes, by clicking the Read Me button. Read the information and address any release-specific requirements.
14. Download the desired software update by clicking the Download button.
The file is downloaded locally.
15. Unzip the downloaded file using an archive manager or by issuing the unzip command.
The file is expanded into the All_Supported_Platforms directory.
16. To upload and apply the update, use the Updating via the BUI or Updating via the CLI procedures in the Maintenance > System online help page.

Connecting the Oracle ZFS Storage ZS3-BA

Oracle Exadata

For each host channel adapter (HCA) on the Oracle ZFS Storage ZS3-BA, connect the HCA's upper port to a corresponding port of the upper InfiniBand gateway switch (leaf switch) on the Oracle Exadata, and the lower port to a corresponding port of the lower leaf switch. Use leaf switch ports: 5B, 6A, 6B, 7A, 7B, or 12A.

Oracle SuperCluster T4-4

Connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T4-4

There are two methods for connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T4-4, depending on the configured environment:

- **Connect directly to the Oracle SuperCluster T4-4 InfiniBand switches:** This option is used if the Oracle ZFS Storage ZS3-BA is the only other appliance or device (other than Oracle SuperCluster T4-4 expansion) that will be connected to the infrastructure.

- **Connect to external InfiniBand leaf switches:** This option is used if more appliances or devices will be connected. In this case, two additional leaf switches are introduced into the InfiniBand network, but are not located within an Oracle SuperCluster T4-4 or Oracle SuperCluster T4-4 expansion rack. A typical scenario for this configuration is when both an Oracle ZFS Storage ZS3-BA and backup application media servers (connected to tape drives) are needed.

Connecting the Oracle ZFS Storage ZS3-BA Directly to the Oracle SuperCluster T4-4 InfiniBand Switches

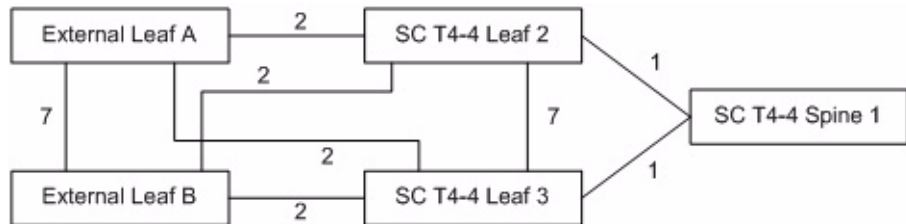
Four ports on each of the Oracle SuperCluster T4-4 InfiniBand leaf switches are used to connect to the Oracle ZFS Storage ZS3-BA. Connect the cables to these ports as follows:

- Oracle ZFS Storage ZS3-BA Head 1
 - PCIe 3 Port 1 to Upper IB Leaf Switch (U24) Port 2A
 - PCIe 3 Port 2 to Lower IB Leaf Switch (U18) Port 2B
 - PCIe 6 Port 1 to Upper IB Leaf Switch (U24) Port 7B
 - PCIe 6 Port 2 to Lower IB Leaf Switch (U18) Port 12A
- Oracle ZFS Storage ZS3-BA Head 2
 - PCIe 3 Port 1 to Lower IB Leaf Switch (U18) Port 2A
 - PCIe 3 Port 2 to Upper IB Leaf Switch (U24) Port 2B
 - PCIe 6 Port 1 to Lower IB Leaf Switch (U18) Port 7B
 - PCIe 6 Port 2 to Upper IB Leaf Switch (U24) Port 12A

Connecting the Oracle SuperCluster T4-4 External InfiniBand Leaf Switches

When external InfiniBand leaf switches are used, the eight ports that were used for directly connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T4-4 infrastructure are used instead for connecting to the external leaf switches. The external leaf switches are connected as shown in [FIGURE 3-1](#), using ports 2A, 2B, 7B, and 12A on each of the Oracle SuperCluster T4-4 leaf switches. The numbers corresponding with the interconnections represent the number of connections between the switches.

FIGURE 3-1 External InfiniBand leaf switches connected to the Oracle SPARC SuperCluster T4-4 InfiniBand leaf switches



The port connections on the external leaf switches are not tightly controlled. Thus, specific external leaf switch port selections are not critical.

Follow these basic guidelines when connecting the Oracle ZFS Storage ZS3-BA to the external leaf switches to ensure maximum throughput and availability:

- Connect the two Port 1 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 1 to two ports on External Leaf A.
- Connect the two Port 2 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 1 to two ports on External Leaf B.
- Connect the two Port 1 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 2 to two ports on External Leaf B.
- Connect the two Port 2 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 2 to two ports on External Leaf A.

Oracle SuperCluster T5-8

Connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T5-8

There are two methods for connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T5-8, depending on the configured environment:

- **Connect directly to the Oracle SuperCluster T5-8 InfiniBand switches:** This option is used if the Oracle ZFS Storage ZS3-BA is the only other appliance or device (other than Oracle SuperCluster T5-8 expansion) that will be connected to the infrastructure.
- **Connect to external InfiniBand leaf switches:** This option is used to increase connection fault tolerance and increase system throughput. It also provides additional ports to attach other systems, such as backup application media servers.

Connecting the Oracle ZFS Storage ZS3-BA Directly to the Oracle SuperCluster T5-8 InfiniBand Switches

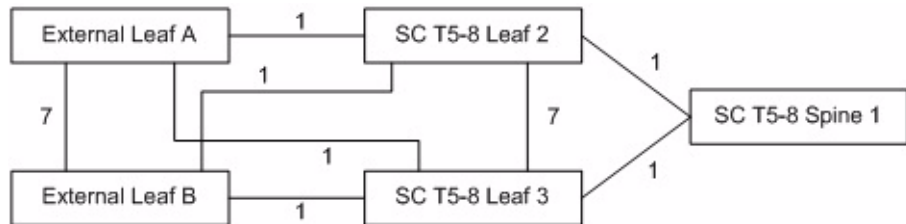
There are two ports available on each of the Oracle SuperCluster T5-8 InfiniBand leaf switches that can be used to connect to the Oracle ZFS Storage ZS3-BA. Connect the cables to these ports as follows:

- Oracle ZFS Storage ZS3-BA Head 1
 - PCIe 3 Port 1 to Upper IB Leaf Switch (U32) Port 2A
 - PCIe 6 Port 1 to Upper IB Leaf Switch (U32) Port 2B
- Oracle ZFS Storage ZS3-BA Head 2
 - PCIe 3 Port 1 to Lower IB Leaf Switch (U26) Port 2A
 - PCIe 6 Port 1 to Lower IB Leaf Switch (U26) Port 2B

Connecting the Oracle ZFS Storage ZS3-BA to Oracle SuperCluster T5-8 External InfiniBand Leaf Switches

When external InfiniBand switches are used, the four ports that were used for directly connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster T5-8 infrastructure are used instead for connecting to the external leaf switches. The external leaf switches are connected as shown in [FIGURE 3-2](#), using ports 2A and 2B on each of the Oracle SuperCluster T5-8 leaf switches. The numbers corresponding with the interconnections represent the number of connections between the switches.

FIGURE 3-2 External InfiniBand leaf switches connected to the Oracle SuperCluster T5-8 InfiniBand leaf switches



The port connections on the external leaf switches are not tightly controlled. Thus, specific external leaf switch port selections are not critical.

Follow these basic guidelines when connecting the Oracle ZFS Storage ZS3-BA to the external leaf switches to ensure maximum throughput and availability:

- Connect the two Port 1 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 1 to two ports on External Leaf A.
- Connect the two Port 2 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 1 to two ports on External Leaf B.
- Connect the two Port 1 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 2 to two ports on External Leaf B.
- Connect the two Port 2 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 2 to two ports on External Leaf A.

Oracle SuperCluster M6-32

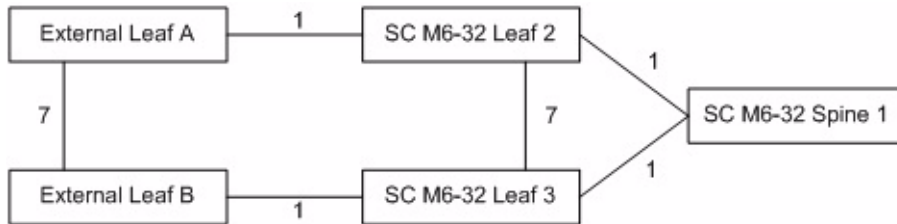
Connecting the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster M6-32

There is only one port on each of the Oracle SuperCluster M6-32 InfiniBand leaf switches available for additional storage connectivity. The two available ports are as follows:

- IB Leaf Switch (U20) Port 12A
- IB Leaf Switch (U22) Port 12A

Therefore, additional external leaf switches are required for connecting an Oracle ZFS Storage ZS3-BA to the infrastructure. The external leaf switches are connected as shown in [FIGURE 3-3](#), using port 12A on each of the Oracle SuperCluster M6-32 leaf switches. The numbers corresponding with the interconnections represent the number of connections between the switches.

FIGURE 3-3 External InfiniBand leaf switches connected to the Oracle SuperCluster M6-32 InfiniBand leaf switches



The port connections on the external leaf switches are not tightly controlled. Thus, specific external leaf switch port selections are not critical.

Follow these basic guidelines when connecting the Oracle ZFS Storage ZS3-BA to the external leaf switches to ensure maximum throughput and availability:

- Connect the two Port 1 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 1 to two ports on External Leaf A.
- Connect the two Port 2 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 1 to two ports on External Leaf B.
- Connect the two Port 1 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 2 to two ports on External Leaf B.
- Connect the two Port 2 InfiniBand HBA ports on Oracle ZFS Storage ZS3-BA Head 2 to two ports on External Leaf A.

Configuring Oracle Exadata Backup

There are two methods for configuring the Oracle ZFS Storage ZS3-BA for the Oracle Exadata:

- Automatic configuration using a utility
- Manual configuration

Both methods enforce a best-practice configuration. The Oracle Engineered Systems Backup Utility is strongly recommended because it reduces human error for complex steps. However, you might opt for manual configuration based on your system or infrastructure requirements.

Automatic Configuration

The Oracle Engineered Systems Backup Utility automates the best-practice configuration setup through a command-line-based wizard. The configuration comprises three steps that are saved individually in a template file, thus allowing later execution for a subsequent step if desired. The three steps can be performed by three different users:

- Step 1 can be performed by the storage administrator to configure the Oracle ZFS Storage ZS3-BA.
- Step 2 can be performed by the Oracle engineered system administrator to configure the Oracle Exadata node.
- Step 3 can be performed by the database administrator to generate the Oracle Recovery Manager (RMAN) run block scripts.

The Oracle Engineered Systems Backup Utility is available on the Oracle Technology Network (<http://www.oracle.com/technetwork/server-storage/sun-unified-storage/downloads/zfssa-plugins-1489830.html>). For information on using the utility, including how to execute a backup, see the user guide and README file, which are bundled with the utility.

Not all configuration has been automated. Configure the network and pools as described in the next section. You do not need to configure the shares or Oracle RMAN.

Manual Configuration

The manual configuration procedure is divided into general configuration guidelines for a quick start, and detailed configuration guidelines for further information.

General Configuration Guidelines

Configuring Networks, Pools, and Shares

The following sections summarize best practices for optimizing Oracle ZFS Storage ZS3-BA network, pool, and share configurations to support backup and restore processing.

Network Configuration

This section describes how to configure the IP network multipathing (IPMP) groups, and how to configure routing in the Oracle ZFS Storage ZS3-BA. The basic network configuration steps are:

1. Connect the Oracle ZFS Storage ZS3-BA to the Oracle Exadata as described in the previous chapter.
2. Configure `ibp0`, `ibp1`, `ibp2`, and `ibp3` with address `0.0.0.0/8` (necessary for IPMP), Connected Mode, and partition key `ffff`. To identify the partition key used by the Oracle Exadata system, run the following command as the root user:

```
# cat /sys/class/net/ib0/pkey
```
3. Configure the active/standby IPMP group over `ibd0` and `ibd3`, with `ibd0` active and `ibd3` standby.
4. Configure the active/standby IPMP group over `ibd1` and `ibd2`, with `ibd2` active and `ibd1` standby.

5. Enable adaptive routing to ensure traffic is load balanced appropriately when multiple IP addresses on the same subnet are owned by the same head. This occurs after a cluster failover.

For customers seeking additional IB connectivity, more IB HCAs can be installed and configured. For details, see the *Oracle ZFS Storage Appliance Installation Guide* (<http://www.oracle.com/goto/ZFSStorage>).

The principles in this section can be applied to a 10 Gb Ethernet implementation by applying the network configuration to the `ixgbe` interfaces instead of the `ibp` interfaces. The 10 Gb Ethernet implementation may be configured as active/active IPMP. If the Oracle ZFS Storage ZS3-BA is on a different subnet than the Oracle Exadata, it may be necessary to create static routes from the Oracle ZFS Storage ZS3-BA to the Oracle Exadata. Consult with your network administrator for details.

Pool Configuration

This section describes design considerations to determine the most appropriate pool configuration for the Oracle ZFS Storage ZS3-BA for Oracle RMAN backup and restore operations based on data protection and performance requirements. The system planner should consider pool protection based on the following guidelines:

- Use parity-based protection for general-purpose and capacity-optimized systems:
 - RAID-Z for protection from single-drive failure on systems subject to random workloads.
 - RAID-Z2 for protection from two-drive failure on systems with streaming workloads only.
- Use mirroring for high-performance with incrementally applied backup.
- Configure pools based on performance requirements:
 - Configure a single pool for management-optimized systems.
 - Configure two pools for performance-optimized systems. Two-pool systems should be configured by using half the drives from each tray.
- Configure log device protection:
 - Stripe log devices for RAID-Z and mirrored pool configurations.
 - Mirror log devices for RAID-Z2 pool configurations.

Share Configuration

The default options for Oracle ZFS Storage ZS3-BA shares provide a good starting point for general-purpose workloads. Oracle ZFS Storage ZS3-BA shares can be optimized for Oracle RMAN backup and restore operations as follows:

- Create a project to store all shares related to backup and recovery of a single database. For a two-pool implementation, create two projects; one for each pool.
- Configure the shares supporting Oracle RMAN backup and restore workloads with the following values:
 - Database record size (`recordsize`): **128kB**
 - Synchronous write bias (`logbias`): **Throughput** (for processing backup sets and image copies) or **Latency** (for incrementally applied backups)
 - Cache device usage (`secondary_cache`): **None** (for backup sets) or **All** (when supporting incrementally applied backups or database clone operations)
 - Data compression (`compression`): **Off** for performance-optimized systems, **LZJB** or **gzip-2** for capacity-optimized systems
 - Number of shares per pool: **1** for management-optimized systems, **2** or **4** for performance-optimized systems

Additional share configuration options, such as higher-level `gzip` compression or replication, can be applied to shares used to support Oracle Exadata backup and restore, as customer requirements mandate.

Customers implementing additional Oracle ZFS Storage ZS3-BA data services should consider implementation-specific testing to verify the implications of deviations from the practices described earlier.

Configuring Oracle RMAN and the Oracle Database Instance

Oracle RMAN is an essential component for protecting the content of Oracle Exadata. Oracle RMAN can be used to create backup sets, image copies, and incrementally updated backups of Oracle Exadata content on Oracle ZFS Storage ZS3-BAs. To optimize performance of Oracle RMAN backups from Oracle Exadata to an Oracle ZFS Storage ZS3-BA, the database administrator should apply the following best practices:

- Load balance Oracle RMAN channels evenly across the nodes of the database machine.
- Load balance Oracle RMAN channels evenly across Oracle ZFS Storage ZS3-BA shares and controllers.

To optimize buffering of the Oracle RMAN channel to the ZFS Storage Appliance, you can tune the values of several hidden instance parameters. For Oracle Database 11g Release 2, the following parameters can be tuned:

- For backup and restore set:
 - `_backup_disk_bufcnt=64`
 - `_backup_disk_bufsz=1048576`

- For image copy backup and restore:
 - `_backup_file_bufcnt=64`
 - `_backup_file_bufsz=1048576`

For additional information about tuning these parameters and tuning equivalent parameters for earlier versions of the Oracle Database software, see Article ID 1072545.1: *RMAN Performance Tuning Using Buffer Memory Parameters* at <http://support.oracle.com>.

Oracle Direct NFS (dNFS) is a high-performance NFS client that delivers exceptional performance for Oracle RMAN backup and restore operations. dNFS should be configured for customers seeking maximum throughput for backup and restore operations.

Detailed Implementation Guidelines

This section provides detailed instructions for implementing an Oracle ZFS Storage ZS3-BA for backup and recovery of the Oracle Exadata.

Configuring the Oracle ZFS Storage ZS3-BA Network

Oracle ZFS Storage ZS3-BA network configuration steps include assigning IP addresses, and optionally IPMP groups, to the physical network interface cards (NICs). For maximum throughput, set the Link Mode to `Connected Mode` for IB interfaces and select `Use Jumbo Frames` for 10 Gb Ethernet. See [FIGURE 4-1](#) for an example of configuring the datalink for an IB interface. See [FIGURE 4-2](#) for an example of configuring the datalink for a 10 Gb Ethernet interface.

Note – All screen shots in the following figures are examples and represent an Oracle ZFS Storage ZS3-BA with four disk shelves.

FIGURE 4-1 Configuring the datalink for an IB interface

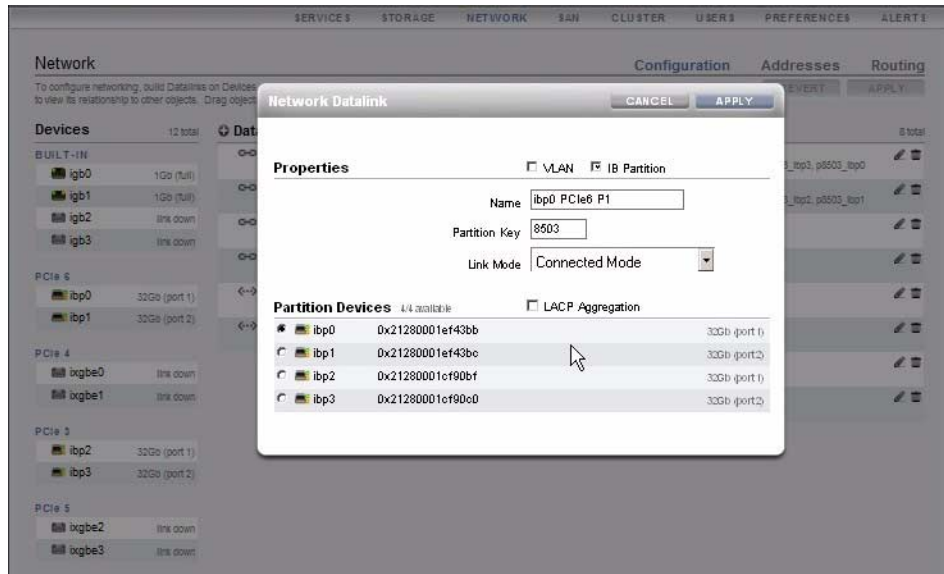
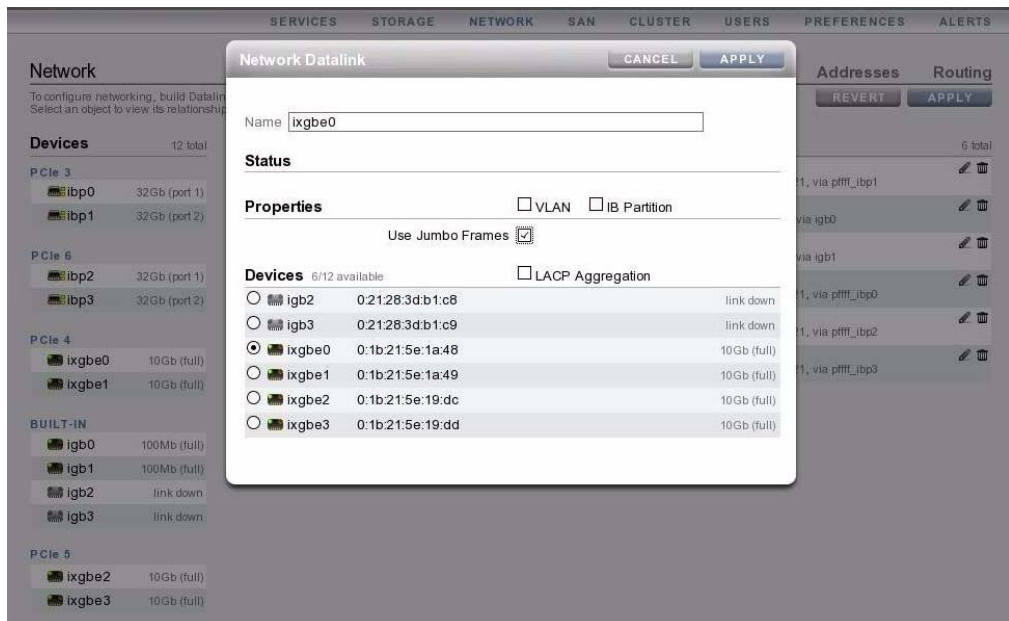


FIGURE 4-2 Configuring the datalink for a 10 Gb Ethernet interface



Configure a minimum of two NICs in each head. Depending on the specific requirements for fault tolerance, IB ports can be configured with fixed IP addresses that can be failed over between cluster controllers, or IPMP, which allows IP addresses to be failed over between ports on a single controller. In practice, the bandwidth of the PCI-2 slot is only able to support the data rate of a single QDR IB port, so a practical IPMP configuration for InfiniBand ports is active/standby access between the ports of a specific card. See [FIGURE 4-3](#), [FIGURE 4-4](#), [FIGURE 4-5](#), and [FIGURE 4-6](#) for examples of configuring these network parameters.

FIGURE 4-3 Completed datalink configuration

The screenshot displays the Oracle Exadata Backup Configuration interface, specifically the **Network** tab. The interface is divided into several sections: **Configuration**, **Addresses**, and **Routing**. The **Configuration** section is active, showing a list of **Devices**, **Datalinks**, and **Interfaces**.

Devices (12 total):

- PCIe 3:**
 - ibp0:** 32Gb (port 1)
 - ibp1:** 32Gb (port 2)
- PCIe 6:**
 - ibp2:** 32Gb (port 1)
 - ibp3:** 32Gb (port 2)
- BUILT-IN:**
 - igb0:** 100Mb (full)
 - igb1:** 100Mb (full)
 - igb2:** link down
 - igb3:** link down
- PCIe 4:**
 - ixgbe0:** 10Gb (full)
 - ixgbe1:** 10Gb (full)
- PCIe 5:**
 - ixgbe2:** 10Gb (full)
 - ixgbe3:** 10Gb (full)

Datalinks (10 total):

- aie-7420a-h1:** via igb0
- aie-7420a-h2:** via igb1
- ibp0:** pkey(ffff), Link Mode(cm), via ibp0
- ibp1:** pkey(ffff), Link Mode(cm), via ibp1
- ibp2:** pkey(ffff), Link Mode(cm), via ibp2
- ibp3:** pkey(ffff), Link Mode(cm), via ibp3
- ixgbe0:** Jumbo frames, via ixgbe0
- ixgbe1:** Jumbo frames, via ixgbe1
- ixgbe2:** Jumbo frames, via ixgbe2
- ixgbe3:** Jumbo frames, via ixgbe3

Interfaces (2 total):

- aie-7420a-h1:** IPv4 static, 10.80.44.17/23, via igb0
- aie-7420a-h2:** IPv4 static, 10.80.44.27/23, via igb1

FIGURE 4-4 Base IPMP interface configured with IB datalink

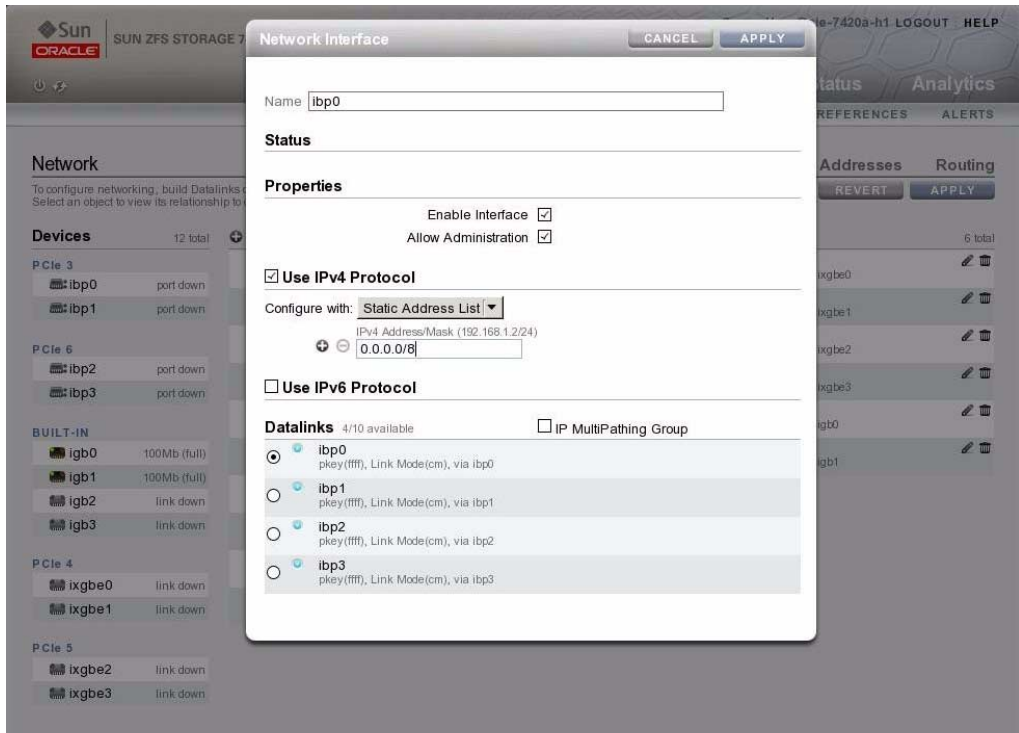


FIGURE 4-5 First active/standby IPMP interface over IB configured

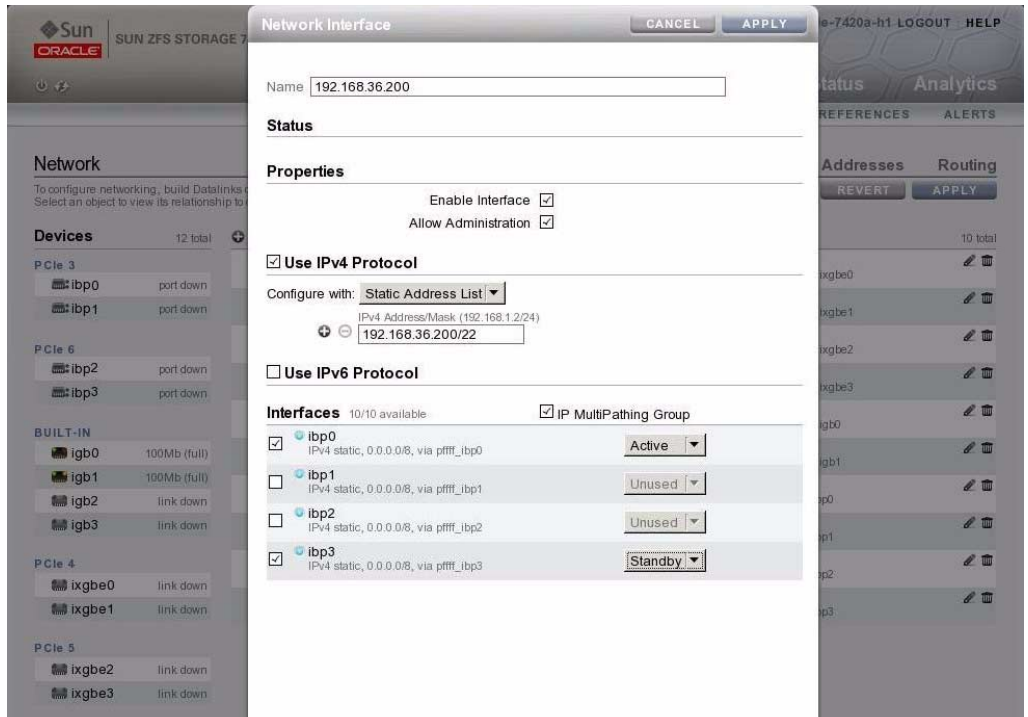
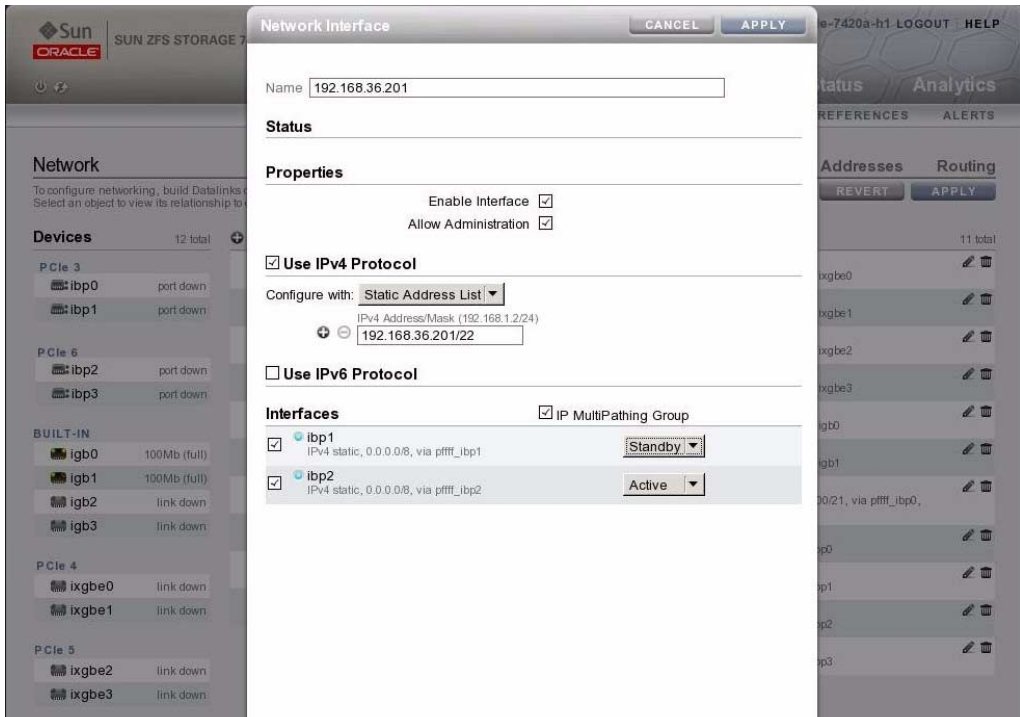


FIGURE 4-6 Second active/standby IPMP interface over IB configured



Configure adaptive routing to allow for data to be returned on the same interface from which it was requested, as shown in [FIGURE 4-7](#).

FIGURE 4-7 Configuring adaptive routing

The screenshot shows the Oracle Exadata Backup configuration interface. The top navigation bar includes tabs for SERVICES, STORAGE, NETWORK, SAN, CLUSTER, USERS, PREFERENCES, and ALERTS. The NETWORK tab is selected, and the sub-tab is 'Routing'. Under 'Properties', the 'Multihoming model' is set to 'Adaptive' (selected with a radio button). Below this, the 'Routing Table Entries' section shows a table with 10 total entries. The table has columns for DESTINATION, GATEWAY, FAMILY, TYPE, and INTERFACE. The entries are as follows:

DESTINATION	GATEWAY	FAMILY	TYPE	INTERFACE
default	10.80.45.254	IPv4	static	10.80.44.27 IPv4 static, 10.80.44.27/23, via igb1
default	10.80.45.254	IPv4	static	10.80.44.17 IPv4 static, 10.80.44.17/23, via igb0
10.80.44.0/23	10.80.44.27	IPv4	system	10.80.44.27 IPv4 static, 10.80.44.27/23, via igb1
10.80.44.0/23	10.80.44.17	IPv4	system	10.80.44.17 IPv4 static, 10.80.44.17/23, via igb0
192.168.32.0/21	192.168.36.201	IPv4	system	192.168.36.201 IPMP, IPv4 static, 192.168.36.201/21, via pffff_ibp1, pffff_ibp2
192.168.32.0/21	192.168.36.200	IPv4	system	192.168.36.200 IPMP, IPv4 static, 192.168.36.200/21, via pffff_ibp0, pffff_ibp3

Note – Depending on how the Oracle ZFS Storage ZS3-BA is installed and configured, the administrator may also need to add a static route to the 1 GB management interface used by the cluster peer.

Configuring the Oracle ZFS Storage ZS3-BA Storage Pool

Pool configuration assigns physical disk drive resources to logical storage pools for backup data storage. To maximize system throughput, configure two equally sized storage pools by assigning half of the physical drives in each drive tray to each storage pool as shown in [FIGURE 4-8](#).

FIGURE 4-8 Storage pool configured based on half of drives in each tray

Configure external storage
ABORT
COMMIT

Allocate and verify storage
Step 1 of 2

Before configuration, you must allocate storage to the pool. SATA storage must be added in whole or half-chassis units; storage in SAS-2 enclosures may be added on a per-device basis. You may leave some storage unallocated. This step will also verify that all devices are present and minimally functional. You can configure storage with affected devices, but they will not be available to the storage pool and cannot be added later without reconfiguring all storage. It is recommended that configuring the affected chassis be deferred until any problems can be repaired.

NAME	MODEL	ALLOCATION	DATA	LOG	CACHE	
aie-7420a-h1	Sun ZFS Storage 7420		-	-	2 (954G)	
0945QCC004	Sun Disk Shelf (SAS-2)		12 (43.7T)	-	-	
1001QCC032	Sun Disk Shelf (SAS-2)		10 (36.4T)	2 (68G)	-	
1001QCC024	Sun Disk Shelf (SAS-2)		12 (43.7T)	-	-	
1001QCC02D	Sun Disk Shelf (SAS-2)		12 (43.7T)	-	-	

Note – The Oracle ZFS Storage ZS3-BA management software presents a warning message about efficiency when two pools with the same RAID protection profile are configured. This message can be safely ignored when configuring for a high-performance Oracle RMAN backup solution.

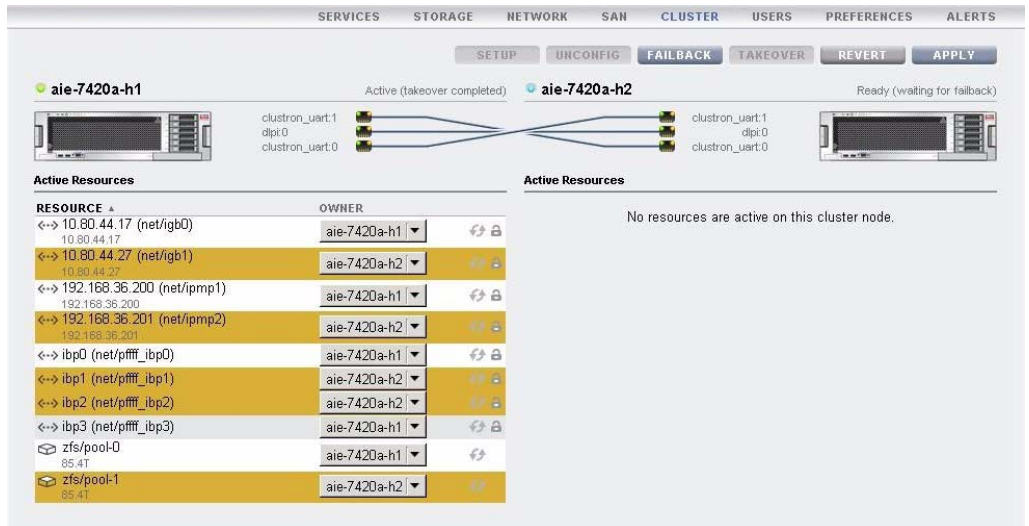
Configuring the Oracle ZFS Storage ZS3-BA Cluster

Cluster configuration assigns physical network and pool resources to specific controllers in the Oracle ZFS Storage ZS3-BA cluster. An important concept in Oracle ZFS Storage ZS3-BA clustering is that a single head owns a specific physical resource at any given time. Consequently, ownership of IP addresses used to access a specific head must be correlated with the pool and network interfaces that are also owned by that controller.

For example, an Oracle ZFS Storage ZS3-BA cluster with two IP addresses on an IPMP group, four IB ports, and two pools, should be configured such that each cluster node owns one of the two IPMP groups, the requisite interface ports to support the group, and one of the two pools. Client access to a specific share on a specific pool must be executed using the IP address associated with the pool.

FIGURE 4-9 shows a cluster configuration on the Oracle ZFS Storage ZS3-BA.

FIGURE 4-9 Cluster configuration for the Oracle ZFS Storage ZS3-BA



Configuring the Oracle ZFS Storage ZS3-BA Share

Share configuration is the process of setting up and tuning NFS mount points for client access. For implementations with two pools (for example, `pool-1` and `pool-2`), create two separate projects named `dbname` in `pool-1` and `pool-2`, respectively.

A project is an Oracle ZFS Storage ZS3-BA entity that provides a higher-level management interface point for a collection of shares. To optimize share management, update the default mount point for shares contained in the project to reference the database name, such as `/export/dbname`. If two projects support the same database, the names of the shares contained in each project must be different. For example, if `dbname` on `pool-1` contains two shares, they could be configured as `backup1` and `backup3`, while the corresponding two shares in `dbname` on `pool-2` could be configured as `backup2` and `backup4`.

For a performance-optimized system, create eight shares for each project. For shares supporting streaming backup and restore operations, set the Database record size property to **128kB** and the Synchronous write bias property to **Throughput**.

FIGURE 4-10 shows a sample share configuration that meets these specifications.

FIGURE 4-10 Share configuration setting appropriate for either a one- or two-pool configuration

The screenshot shows the Oracle Exadata configuration interface for a share named 'dbname'. The interface is divided into several sections:

- Top Navigation:** SHARES, PROJECTS, SCHEMA. The 'PROJECTS' tab is active, showing a list of projects with 'dbname' selected.
- Share Configuration:**
 - General:** Shows the share name 'dbname' and its path 'pool-0/localhost/dbname'. Buttons for 'REVERT' and 'APPLY' are present.
 - Space Usage:**
 - DATA:** Quota and Reservation are both set to 0 GB.
 - USERS & GROUPS:** A dropdown menu for 'User or Group' is set to 'none'. A 'Show All' link is available.
 - Inherited Properties:**
 - Mountpoint: /export/dbname
 - Read only: ☐
 - Update access time on read: ☐
 - Non-blocking mandatory locking: ☐
 - Data deduplication (warning): ☐
 - Data compression: Off
 - Checksum: Fletcher4 (Standard)
 - Cache device usage: All data and metadata
 - Synchronous write bias: Throughput
 - Database record size: 128k
 - Additional replication: Normal (Single Copy)
 - Virus scan: ☐
 - Prevent destruction: ☐
 - Default Settings:**
 - FILESYSTEMS:**
 - User: nobody
 - Group: other
 - Permissions: R W X (User), R W X (Group), R W X (Other)
 - LUNS:**
 - Volume size: 0 GB
 - Thin provisioned: ☐
 - Volume block size: 8k

Root access can be granted for all database nodes in an Oracle Exadata rack by specifying a specific NFS access rule to allow root-access for any server on the Oracle Exadata private network, for example, 192.168.36.0/22. [FIGURE 4-11](#) shows an example of NFS root-access exceptions configured for a specific project.

FIGURE 4-11 Granting root access for all database nodes in an Oracle Exadata rack

The screenshot shows the Oracle Exadata configuration interface for a share named 'dbname'. The interface is divided into several tabs: SHARES, PROJECTS, and SCHEMA. The 'SHARES' tab is active, and the 'Access' sub-tab is selected. The share is located at 'pool-0/localhost/dbname' and is of type 'NFS'. The path is 'aie-7420a-h1/export/dbname'. The share mode is set to 'Read/write'. There are checkboxes for 'Disable setuid/setgid file creation' and 'Prevent clients from mounting subdirectories', both of which are unchecked. The 'Anonymous user mapping' is set to 'nobody'. The 'Character encoding' is set to 'default'. The 'Security mode' is set to 'Default (AUTH_SYS)'. Below these settings is a table for 'NFS Exceptions'. The table has columns for TYPE, ENTITY, ACCESS MODE, CHARSET, and ROOT ACCESS. There is one exception entry with TYPE 'Network', ENTITY '192.168.36.0/22', ACCESS MODE 'Read/write', CHARSET 'default', and ROOT ACCESS checked. Below the table are sections for 'SMB', 'HTTP', 'FTP', and 'SFTP'. Each section has a 'Share mode' dropdown menu. For SMB, the 'Resource Name' is 'off', 'Enable Access-based Enumeration' is unchecked, and 'Is a DFS Namespace' is 'No'. For HTTP, the 'Share mode' is 'None'. For FTP, the 'Share mode' is 'None'. For SFTP, the 'Share mode' is 'None'.

TYPE	ENTITY	ACCESS MODE	CHARSET	ROOT ACCESS
Network	192.168.36.0/22	Read/write	default	<input checked="" type="checkbox"/>

Configuring the Oracle ZFS Storage ZS3-BA DTrace Analytics

The Oracle ZFS Storage ZS3-BA includes a comprehensive performance analysis tool called DTrace Analytics. DTrace Analytics is a framework that monitors important subsystem performance accounting statistics. A subset of the available accounting statistics should be monitored to provide comprehensive data on the effectiveness and performance of Oracle RMAN backup and restore workloads.

The following Analytics are available when advanced analytics are configured on the Oracle ZFS Storage ZS3-BA (Configuration > Preferences > Enable Advanced Analytics):

- **CPU:** Percent utilization broken down by CPU mode
- **Disk:** Average number of I/O operations broken down by state of operation

- *Disk*: I/O bytes per second broken down by type of operation
- *Disk*: I/O operations per second broken down by latency
- *Disk*: Disks with utilization of at least 95 percent broken down by disk
- *Network*: Interface bytes per second broken down by direction
- *Network*: Interface bytes per second broken down by interface
- *Protocol*: NFSv3 operations per second broken down by size
- *Protocol*: NFSv3 operations per second broken down by type of operation
- *Protocol*: NFSv3 operations per second of type read broken down by latency
- *Protocol*: NFSv3 operations per second of type write broken down by latency
- *Protocol*: NFSv3 operations per second of type read broken down by size
- *Protocol*: NFSv3 operations per second of type write broken down by size

Implementing these accounting statistics helps end-users gain a quantitative understanding of the instantaneous and historical resource consumption and quality of service (QoS) for their specific implementation.

Configuring the Client NFS Mount

When configuring the Oracle ZFS Storage ZS3-BA, any server that accesses the appliance, including Oracle Exadata servers, is considered a client. Configuring the client NFS mount includes creating the target directory structure for access to the Oracle ZFS Storage ZS3-BA as well as the specific NFS mount options necessary for optimal system performance. Mount options for Linux clients are:

```
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,vers=3,timeo=600
```

For detailed configuration steps, see [“Configuration Details for Client NFS Mount and Oracle dNFS” on page 79](#).

Note – Implementations using Oracle Database version 11.2.0.1 may run into the bug 9244583, ORA-27054 when running Oracle RMAN over NFS. The patch for this bug and workarounds are documented in the My Oracle Support document ORA-27054 WHEN RUNNING RMAN WITH NFS IN 11.2 (WORKS FINE ON 10.2 AND 11.1) [ID 1076405.1].

Tuning the Linux Network and Kernel

Depending on the specific Linux installation, the NFS client software and necessary supporting software subsystems may or may not be enabled. Two Linux services required to run NFS are portmap and nfslock. The services can be configured to run after reboot using the `chkconfig` command and enabled dynamically using the `service` command as follows:

```
# chkconfig portmap on
# service portmap start
# chkconfig nfslock on
# service nfslock start
```

To ensure the portmap service has access to the `/etc/hosts.allow` and `/etc/hosts.deny` files, open up group and world read permissions on these files after verifying with local system administration officials that read permissions may be granted for these files:

```
# ls -l /etc/host*
-rw-r--r-- 1 root root  17 Jul 23  2000 /etc/host.conf
-rw-r--r-- 1 root root 1394 Mar  4 10:36 /etc/hosts
-rw----- 1 root root  161 Jan 12  2000 /etc/hosts.allow
-rw-r--r-- 1 root root  147 Mar  3 14:03 /etc/hosts.backupbyExadata
-rw----- 1 root root  347 Jan 12  2000 /etc/hosts.deny
-rw-r--r-- 1 root root  273 Mar  3 14:03 /etc/hosts.orig

# dcli -l root -g /home/oracle/dbs_group chmod 640 /etc/hosts.allow
# dcli -l root -g /home/oracle/dbs_group chmod 640 /etc/hosts.deny
# dcli -l root -g /home/oracle/dbs_group chown root:rpc
/etc/hosts.allow
# dcli -l root -g /home/oracle/dbs_group chown root:rpc
/etc/hosts.deny

# ls -l /etc/host*
-rw-r--r-- 1 root root  17 Jul 23  2000 /etc/host.conf
-rw-r--r-- 1 root root 1394 Mar  4 10:36 /etc/hosts
-rw-r----- 1 root rpc  161 Jan 12  2000 /etc/hosts.allow
-rw-r--r-- 1 root root  147 Mar  3 14:03 /etc/hosts.backupbyExadata
-rw-r----- 1 root rpc  347 Jan 12  2000 /etc/hosts.deny
-rw-r--r-- 1 root root  273 Mar  3 14:03 /etc/hosts.orig
```

For Oracle Exadata, the Linux service `cpuspeed` is disabled by default, which optimizes throughput for some network devices. In a general Linux implementation, `cpuspeed` may be set to enable by default, which can reduce NFS throughput over

10 Gb Ethernet. If this service is not being used, or its use is less valuable than maximizing NFS performance over 10 Gb Ethernet, the service can be manually disabled after boot or dynamically disabled with the `chkconfig` and `service` commands as follows:

```
# chkconfig cpuspeed off
# service cpuspeed stop
```

Further client, operating system, network, and kernel tuning may be needed, including software updates, to maximize device driver, networking, and kernel throughput related to network I/O processing. These tuning procedures are system-specific and beyond the scope of this paper. Consult with your operating system and NIC vendors for evaluation and implementation details.

Configuring Oracle Direct NFS (dNFS)

A complete description of Direct NFS (dNFS) configuration is available for each specific release of the Oracle Database software from <http://support.oracle.com>.

For detailed configuration steps, see “Configuration Details for Client NFS Mount and Oracle dNFS” on page 79.

Note – Prior to configuring dNFS, apply Oracle Database patch 8808984 to ensure optimal dNFS operation. Patch 8808984 is available from <http://support.oracle.com>, and is included in Oracle Exadata Database 11.2.0.1 (BP 8).

For Oracle Exadata X2-8 and X3-8 configurations, use static routes as described in *How to Setup Direct NFS client multipaths in same subnet* (Note 822481.1) available at <http://support.oracle.com>.

A summary of how to configure dNFS is as follows:

1. Shut down the running instance of the Oracle Database software.
2. Enable dNFS using one of the options below:
 - For *version 11.2.0.2 or greater* of the Oracle Database software, enter:

```
$ make -f \
    $ORACLE_HOME/rdbms/lib/ins_rdbms.mk
dnfs_on
```
 - For a *version prior to 11.2.0.2*, enter:

```
$ ln -sf \
```

```
$ORACLE_HOME/lib/libnfsodm11.so \  
$ORACLE_HOME/lib/libodm11.so
```

3. Update the `orantfstab (/etc/orantfstab)` file with entries showing the channels and shares accessed on the Oracle ZFS Storage ZS3-BA in cases where multiple IP addresses are used to access a single share. The following example shows how to access the backup share on `aie-7420a-h1` over two separate IP addresses, 192.168.36.200 and 192.168.36.201.

```
server: aie-test-1-71-ib-data  
local: 192.168.36.100 path: 192.168.36.200  
local: 192.168.36.101 path: 192.168.36.202  
local: 192.168.36.102 path: 192.168.36.204  
local: 192.168.36.103 path: 192.168.36.206  
dontroute  
export: /export/qs/backup1 mount: /zfssa/qs/backup1  
export: /export/qs/backup3 mount: /zfssa/qs/backup3  
export: /export/qs/backup5 mount: /zfssa/qs/backup5  
export: /export/qs/backup7 mount: /zfssa/qs/backup7  
server: aie-test-1-72-ib-data  
local: 192.168.36.100 path: 192.168.36.201  
local: 192.168.36.101 path: 192.168.36.203  
local: 192.168.36.102 path: 192.168.36.205  
local: 192.168.36.103 path: 192.168.36.207  
dontroute  
export: /export/qs/backup2 mount: /zfssa/qs/backup2  
export: /export/qs/backup4 mount: /zfssa/qs/backup4  
export: /export/qs/backup6 mount: /zfssa/qs/backup6  
export: /export/qs/backup8 mount: /zfssa/qs/backup8
```

4. Restart the Oracle Database software instance.

Tuning the Oracle Database Instance for Oracle RMAN Backup and Restore

Optimizing high-bandwidth backup and restore operations using Oracle RMAN and the Oracle ZFS Storage Appliance requires adjusting the instance parameters that control I/O buffering. For information about how to tune these parameters on different versions of the Oracle Database software, see Article ID 1072545.1: *RMAN Performance Tuning Using Buffer Memory Parameters*) at <http://support.oracle.com>.

For Oracle Exadata, tuning the following four parameters should be considered:

- `_backup_disk_bufcnt` - Number of buffers used to process backup sets
- `_backup_disk_bufsz` - Size of the buffers used to process backup sets

- `_backup_file_bufcnt` - Number of buffers used to process image copies
- `_backup_file_bufsz` - Size of the buffers used to process image copies

For backup and restore operations on backup sets and image copies, set the number of buffers to 64 and the buffer size to 1 MB:

```
SQL> alter system set "_backup_disk_bufcnt"=64;
SQL> alter system set "_backup_file_bufcnt"=64;
SQL> alter system set "_backup_disk_bufsz"=1048576;
SQL> alter system set "_backup_file_bufsz"=1048576;
```

These commands may be configured persistently by adding them to the `SPFILE`, or they may be set dynamically in the Oracle RMAN run block used to execute the backup or restore operations.

The following code fragments show how to dynamically tune the buffer sizes and counts for backup and restore operations.

- Backup set backup:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    allocate channel...
...
    backup as backupset database;
}
```

- Backup set restore:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    allocate channel...
...
    restore database;
}
```

- Image copy backup:

```
run
{
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...
...
    backup as copy database;
}
```


■ Image copy restore:

```
run
{
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...
    ...
    restore database;
}
```

Performing an incrementally applied backup requires reading an incremental backup set and writing to an image copy. To tune buffers for incrementally applied backups, run the following:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...
    ...
    recover copy of database;
}
```

Creating Dedicated Services for Oracle RMAN Operations

Eight services dedicated to Oracle RMAN processing can be configured to optimize management of load balancing, high availability, and upgrades. These services can be evenly load balanced over all the nodes of an Oracle Exadata system. Availability and performance can be optimized by configuring the services to run on a preferred instance while preparing them to fail over to any instance in the cluster. If these services are configured, upgrading a one-quarter or one-half rack Oracle Exadata system does not require changing the connect string of the Oracle RMAN run block.

The `srvctl` utility is used to install services for Oracle RMAN processing. The following code fragment shows how to create eight services evenly distributed over a four-node cluster that are set up to fail over to any other node in the cluster. In this example, the services are installed for a database named `dbname` and are named `dbname_bkup[1-8]`.

```
srvctl add service -d dbname -r dbname1 -a dbname2,dbname3,dbname4 \
                -s dbname_bkup1
srvctl start service -d dbname -s dbname_bkup1
```

```

srvctl add service -d dbname -r dbname2 -a dbname1,dbname3,dbname4 \
-s dbname_bkup2
srvctl start service -d dbname -s dbname_bkup2
srvctl add service -d dbname -r dbname3 -a dbname1,dbname2,dbname4 \
-s dbname_bkup3
srvctl start service -d dbname -s dbname_bkup3
srvctl add service -d dbname -r dbname4 -a dbname1,dbname2,dbname3 \
-s dbname_bkup4
srvctl start service -d dbname -s dbname_bkup4
srvctl add service -d dbname -r dbname1 -a dbname2,dbname3,dbname4 \
-s dbname_bkup5
srvctl start service -d dbname -s dbname_bkup5
srvctl add service -d dbname -r dbname2 -a dbname1,dbname3,dbname4 \
-s dbname_bkup6
srvctl start service -d dbname -s dbname_bkup6
srvctl add service -d dbname -r dbname3 -a dbname1,dbname2,dbname4 \
-s dbname_bkup7
srvctl start service -d dbname -s dbname_bkup7
srvctl add service -d dbname -r dbname4 -a dbname1,dbname2,dbname3 \
-s dbname_bkup8
srvctl start service -d dbname -s dbname_bkup8

```

Configuring Oracle RMAN

Configuring Oracle RMAN channel and parallelism includes specifying the file system targets for the Oracle RMAN backup channels and the total number of channels used for backup and restore operations. Performance benefits can be realized by configuring 16 Oracle RMAN channels spanning the available Oracle ZFS Storage ZS3-BA shares. Configure Oracle RMAN channels such that they are evenly distributed over the Oracle Database instances and nodes in the RAC cluster and evenly distributed over the shares exported from the Oracle ZFS Storage ZS3-BA.

The following code fragments show sample Oracle RMAN run blocks for performing backup and restore operations for backup sets and image copies as well as applying incremental merges to image copies. The sample code is based on the following database configuration:

- Database name: dbname
- SYSDBA login: sys/welcome
- Scan address: ad01-scan
- Service names for the backup: dbname_bkup[1-8]

The Oracle ZFS Storage ZS3-BA can be configured in a one-pool configuration in which the appliance exports eight shares used as eight mount points.

The Oracle RMAN run blocks for backup and restore using backup sets and image copies are shown in the examples in the sections below. In these examples, the mount points for the four-share configuration are accessed as /zfssa/dbname/backup1 through /zfssa/dbname/backup4. Also, the examples are for a configuration in which the Oracle ZFS Storage ZS3-BA exports four shares used as four mount points for 16 Oracle RMAN channels.

Backup set level 0 backup:

```
run
{
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3' format '/zfssa/dbname/backup3/%U';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4' format '/zfssa/dbname/backup4/%U';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5' format '/zfssa/dbname/backup1/%U';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6' format '/zfssa/dbname/backup2/%U';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7' format '/zfssa/dbname/backup3/%U';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8' format '/zfssa/dbname/backup4/%U';

  allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3' format '/zfssa/dbname/backup4/%U';
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4' format '/zfssa/dbname/backup3/%U';
  allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5' format '/zfssa/dbname/backup2/%U';
  allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6' format '/zfssa/dbname/backup1/%U';
```

```

allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7' format '/zfssa/dbname/backup4/%U';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8' format '/zfssa/dbname/backup3/%U';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

backup as backupset incremental level 0 section size 32g database
tag 'FULLBACKUPSET_L0' plus archivelog tag 'FULLBACKUPSET_L0';
}

```

Backup set level 1 backup:

```

run
{
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3' format '/zfssa/dbname/backup3/%U';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4' format '/zfssa/dbname/backup4/%U';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5' format '/zfssa/dbname/backup1/%U';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6' format '/zfssa/dbname/backup2/%U';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7' format '/zfssa/dbname/backup3/%U';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8' format '/zfssa/dbname/backup4/%U';

  allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3' format '/zfssa/dbname/backup4/%U';
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4' format '/zfssa/dbname/backup3/%U';
}

```

```

allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5' format '/zfssa/dbname/backup2/%U';
allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6' format '/zfssa/dbname/backup1/%U';
allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7' format '/zfssa/dbname/backup4/%U';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8' format '/zfssa/dbname/backup3/%U';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

backup as backupset incremental level 1 database tag
'FULLBACKUPSET_L1' plus archivelog tag 'FULLBACKUPSET_L1';
}

```

Image copy backup:

```

run
{
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3' format '/zfssa/dbname/backup3/%U';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4' format '/zfssa/dbname/backup4/%U';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5' format '/zfssa/dbname/backup1/%U';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6' format '/zfssa/dbname/backup2/%U';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7' format '/zfssa/dbname/backup3/%U';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8' format '/zfssa/dbname/backup4/%U';

  allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
}

```

```

allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3' format '/zfssa/dbname/backup4/%U';
allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4' format '/zfssa/dbname/backup3/%U';
allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5' format '/zfssa/dbname/backup2/%U';
allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6' format '/zfssa/dbname/backup1/%U';
allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7' format '/zfssa/dbname/backup4/%U';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8' format '/zfssa/dbname/backup3/%U';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

backup incremental level 1 for recover of copy with tag 'IMAGECOPY'
database;
}

```

Incremental merge to image copy:

```

run
{
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8';
}

```

```

allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3';
allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4';
allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5';
allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6';
allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

recover copy of database with tag 'IMAGECOPY';
}

```

Restore validate:

```

run
{
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6';
}

```

```

allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7';
allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8';

allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup3';
allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup4';
allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup5';
allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup6';
allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup7';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup8';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

restore validate database;
}

```


Configuring Oracle SuperCluster Backup

There are two methods for configuring the Oracle ZFS Storage ZS3-BA for the Oracle SuperCluster:

- Automatic configuration using a utility
- Manual configuration

Both methods enforce a best-practice configuration. The Oracle Engineered Systems Backup Utility is strongly recommended because it reduces human error for complex steps. However, you might opt for manual configuration based on your system or infrastructure requirements.

Automatic Configuration

The Oracle Engineered Systems Backup Utility automates the best-practice configuration setup through a command-line-based wizard. The configuration comprises three steps that are saved individually in a template file, thus allowing later execution for a subsequent step if desired. The three steps can be performed by three different users:

- Step 1 can be performed by the storage administrator to configure the Oracle ZFS Storage ZS3-BA.
- Step 2 can be performed by the Oracle engineered system administrator to configure the Oracle SuperCluster node.
- Step 3 can be performed by the database administrator to generate the Oracle Recovery Manager (RMAN) run block scripts.

The Oracle Engineered Systems Backup Utility is available on the Oracle Technology Network (<http://www.oracle.com/technetwork/server-storage/sun-unified-storage/downloads/zfssa-plugins-1489830.html>). For information on using the utility, including how to execute a backup, see the user guide and README file, which are bundled with the utility.

General Implementation Tasks

The following sections summarize best practices for optimizing the Oracle ZFS Storage ZS3-BA configuration for the Oracle SuperCluster 11gR2.

- “Configuring the Oracle ZFS Storage ZS3-BA InfiniBand Datalinks” on page 51
- “Configuring the Oracle SuperCluster InfiniBand Switches to Add the Oracle ZFS Storage ZS3-BA” on page 52
- “Configuring Oracle ZFS Storage ZS3-BA Networking for Single IP Connection” on page 55
- “Configuring Oracle ZFS Storage ZS3-BA Networking for an Active-Active Configuration” on page 57
- “Configuring the Oracle ZFS Storage ZS3-BA Storage Pool” on page 62
- “Configuring the Oracle ZFS Storage ZS3-BA Shares” on page 62
- “Configuring the Oracle ZFS Storage ZS3-BA DTrace Analytics” on page 65
- “Configuring the Client NFS Mount” on page 66
- “Tuning the Solaris 11 Network and Kernel” on page 66
- “Configuring Oracle Direct NFS (dNFS)” on page 67
- “Tuning the Oracle Database Instance for Oracle RMAN Backup and Restore” on page 68
- “Creating Dedicated Services for Oracle RMAN Operations” on page 70
- “Configuring Oracle RMAN” on page 71

Detailed Implementation Tasks

The following sections provide detailed information for performing the best-practices tasks for optimizing the Oracle ZFS Storage ZS3-BA configuration.

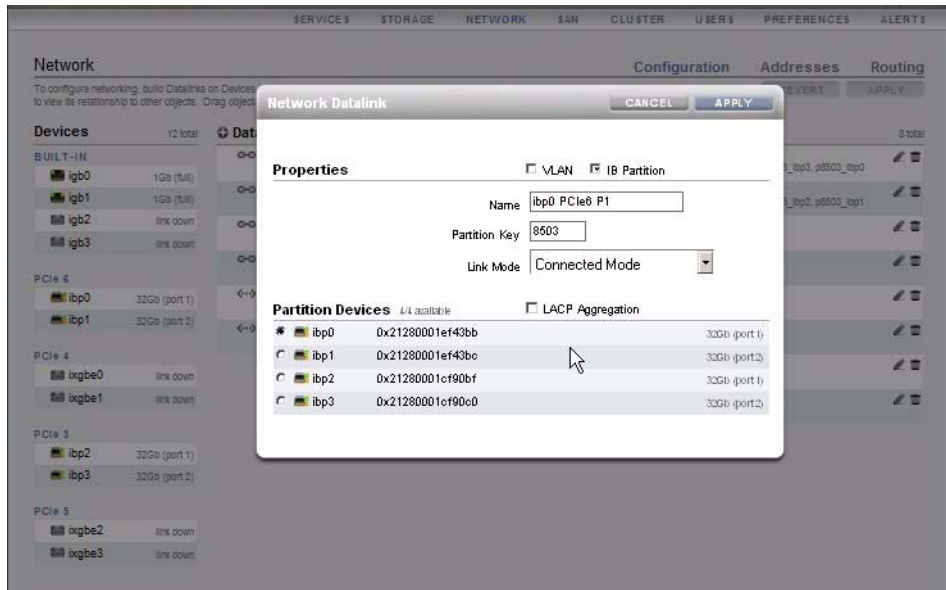
All screen shots, print outs, and guidelines in the following sections are examples and represent an Oracle ZFS Storage ZS3-BA with two controllers (heads) and four disk shelves.

Configuring the Oracle ZFS Storage ZS3-BA InfiniBand Datalinks

Follow the steps in this section to configure each Oracle ZFS Storage ZS3-BA InfiniBand connection. The eight GUIDs for the InfiniBand HBA ports that are recorded during this procedure are used to configure the Oracle SuperCluster InfiniBand switches in the next procedure.

1. Connect the Oracle ZFS Storage ZS3-BA to the Oracle SuperCluster as described in [“Connecting the Oracle ZFS Storage ZS3-BA” on page 15](#).
2. Log on to the Browser User Interface (BUI) of Head 1 and navigate to **Configuration > Network**.
3. Click the plus icon next to **Datalinks**. The Network Datalink dialogue box opens.
4. Complete the dialogue box as follows (refer to [FIGURE 5-1](#)):
 - a. Check the `IB Partition` box.
 - b. Enter a meaningful name for the datalink name.
 - c. Set the `Partition Key` to **8503**.
 - d. Select **Connected Mode** for the `Link Mode`.
 - e. Do not check the `LACP Aggregation` box.
 - f. Select `Partition Device` **ibp0**.
 - g. Record the GUID number (for example, 21280001ef43bb) and click **Apply**.

FIGURE 5-1 Network Interface dialogue box



5. Repeat steps 3 and 4 for each remaining InfiniBand interface (**ibp1**, **ibp2**, and **ibp3**).
6. Repeat steps 2 through 5 for Head 2.

Configuring the Oracle SuperCluster InfiniBand Switches to Add the Oracle ZFS Storage ZS3-BA

In this procedure, the GUIDs of the Oracle ZFS Storage ZS3-BA Infiniband HBA ports are added to the existing Oracle SuperCluster InfiniBand configuration. By adding these ports and using a partition key of 8503, communication between the two devices can occur.

1. Log on to the Oracle SuperCluster InfiniBand spine switch as root.

By default, the spine switch is given a hostname of <sscid>sw-ib1, where <sscid> is the prefix name given to the entire Oracle SuperCluster system. In the following example, the <sscid> is **aiessc**.

```
login as: root
root@aiesscsw-ib1's password:
Last login: Tue Sep 25 08:19:01 2013 from dhcp-brm-b15-204-3e-
east-10-135-75-254.usdhcp.oraclecorp.com
```

2. Enter the command `enablesm` to verify that the switch is running Subnet Manager (or this command will start Subnet Manager).

```
[root@aiesscsw-ib1 ~]# enablesm
opensm (pid 15906) is already running...
Starting partitiond daemon
/usr/local/util/partitiond is already running
(You may also perform a 'restart' if wanted)
```

3. Enter the command `getmaster` to verify that this is the master switch of the configuration. If the master switch is not running on the spine switch, log out and log in to the designated master switch for the remainder of this procedure.

```
[root@aiesscsw-ib1 ~]# getmaster
Local SM enabled and running
20130913 10:16:51 Master SubnetManager on sm lid 13 sm guid
0x2128e8ac27a0a0 : SUN DCS 36P QDR aiesscsw-ib1.us.oracle.com
[root@aiesscsw-ib1 ~]#
```

4. Back up the switch configuration according the documented backup procedures (http://docs.oracle.com/cd/E26698_01/index.html).

5. Enter the command `smpartition list active` to verify that partition key 0x0503 is assigned to partition name "sto" (**sto = 0x0503**).

The partition key was set to 8503 on the Oracle ZFS Storage ZS3-BA datalinks, but the InfiniBand switch reports 0503. This is intentional because the InfiniBand protocol reserves the most significant bit (0x8000) of the hexadecimal partition key (pkey) for its own use. Therefore, pkeys 0x8503 and 0x0503 are the same.

```
[root@aiesscsw-ib1 ~]# smpartition list active
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 11
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
ic1s10 = 0x0501, ipoib, defmember=full:
0x0021280001ef30f7,
0x0021280001ef33bf,
0x0021280001ef30b7,
0x0021280001ef314b;
ic2s10 = 0x0502, ipoib, defmember=full:
0x0021280001ef30f8,
0x0021280001ef33c0,
0x0021280001ef30b8,
0x0021280001ef314c;
sto = 0x0503, ipoib, defmember=full:
```

```

0x0021280001ef43f8,
0x0021280001ef43b7,
0x0021280001cf90c0,
0x0021280001ef43bb,
... more ...

```

6. Add the Oracle ZFS Storage ZS3-BA to the InfiniBand configuration:

- a. Enter the command `smpartition start` to start a reconfiguration session.
`smpartition start`

```
[root@aiaesscsw-ib1 ~]# smpartition start
```

- b. Enter the command `smpartition add` to add the eight new GUIDs to the configuration.

```
# smpartition add -n sto -port <GUID1> <GUID2> <GUID3> ...
<GUID8>
```

```
[root@aiaesscsw-ib1 ~]# smpartition add -n sto -port
21280001ef43bb 21280001ef43bc 21280001cf90bf 21280001cf90c0
21280001ef43f7 21280001ef43f8 21280001ef43b7 21280001ef43b8
```

- c. Enter the command `smpartition list modified` to verify the new GUIDs have been added correctly.

```
# smpartition list modified
```

```
[root@aiaesscsw-ib1 ~]# smpartition list modified
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 11
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=
full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
ic1s10 = 0x0501,ipoib,defmember=full:
0x0021280001ef30f7,
0x0021280001ef33bf,
0x0021280001ef30b7,
0x0021280001ef314b;
ic2s10 = 0x0502,ipoib,defmember=full:
0x0021280001ef30f8,
0x0021280001ef33c0,
0x0021280001ef30b8,
0x0021280001ef314c;
sto = 0x0503,ipoib,defmember=full:
0x0021280001ef43f8,
0x0021280001ef43b7,
0x0021280001cf90c0,
0x0021280001ef43bb,
```

```

0x0021280001ef43bc,
0x0021280001cf90bf,
0x0021280001ef43b8,
0x0021280001ef43f7,
0x0021280001ef3048,
0x0021280001ef30af,
0x0021280001ef30f8,
0x0021280001ef30f7,
0x0021280001ef33c0,
0x0021280001ef33bf,
0x0021280001ef30cc,
0x0021280001ef342b,
0x0021280001ef30b8,
0x0021280001ef30b7,
0x0021280001ef314c,
0x0021280001ef314b,
0x0021280001efec65,
0x0021280001efec66,
0x0021280001efecb1,
0x0021280001efecb2;

```

- d. Enter the command `smpartition commit` to apply the new configuration and propagate configuration changes to all InfiniBand switches in the configuration.

```
# smpartition commit
```

```
[root@aiaesscs-w-ib1 ~]# smpartition commit
[root@aiaesscs-w-ib1 ~]#
```

7. Log off the InfiniBand switch.
8. Back up the InfiniBand configuration according to the documented backup procedures (http://docs.oracle.com/cd/E26698_01/index.html).

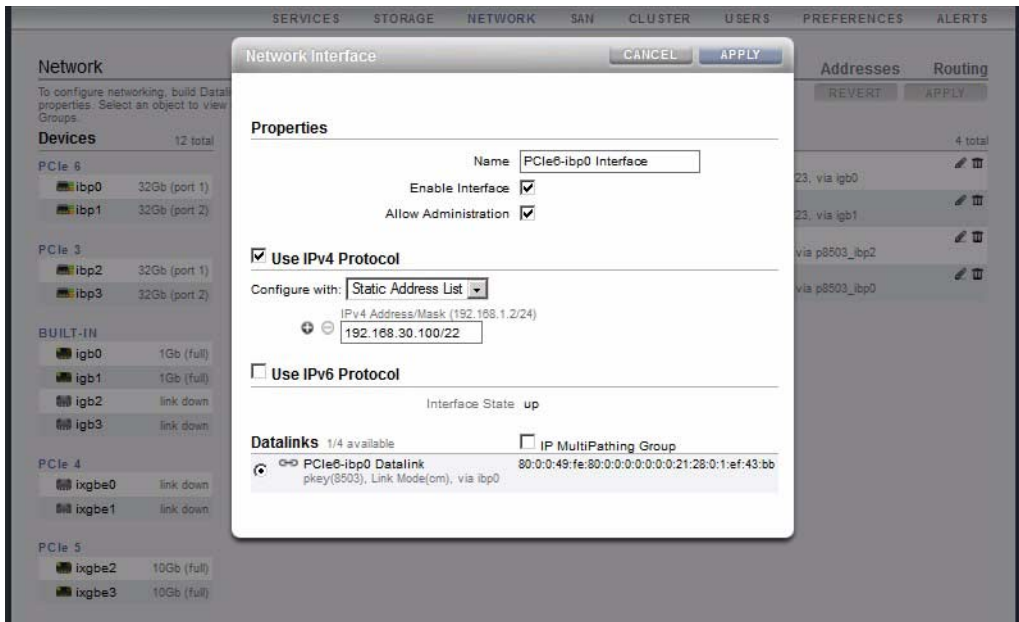
Configuring Oracle ZFS Storage ZS3-BA Networking for Single IP Connection

This configuration is only for an Oracle SuperCluster T5-8 with no external leaf switches. For best failover and performance, use the Active-Active Configuration (next section) for all other configurations.

Configure the Oracle ZFS Storage ZS3-BA InfiniBand ports for network connectivity and simple cluster failover by using the following procedure to configure Port 1 with the desired IP address.

1. Log on to the BUI of Head 1 and navigate to **Configuration > Network**.
2. Click the plus icon next to **Interfaces**. The Network Interface dialogue box opens.
3. Complete the dialogue box as follows (refer to [FIGURE 5-2](#)):
 - a. Enter a meaningful name for the network interface.
 - b. Verify that **Enable Interface** is checked.
 - c. Verify that **Allow Administration** is checked.
 - d. Verify that **Use IPv4 Protocol** is checked.
 - e. Verify that the **Configure with** menu selection is **Static Address List**.
 - f. In the box below that, enter the desired IP address with the appropriate netmask.
 - g. Verify that **Use IPv6 Protocol** is not checked.
 - h. Select the datalink for **ibp0** and click **Apply**.
4. Repeat steps 1 through 3 on Head 2 using **ibp2** as the datalink.

FIGURE 5-2 Configuring an IP address for an InfiniBand datalink

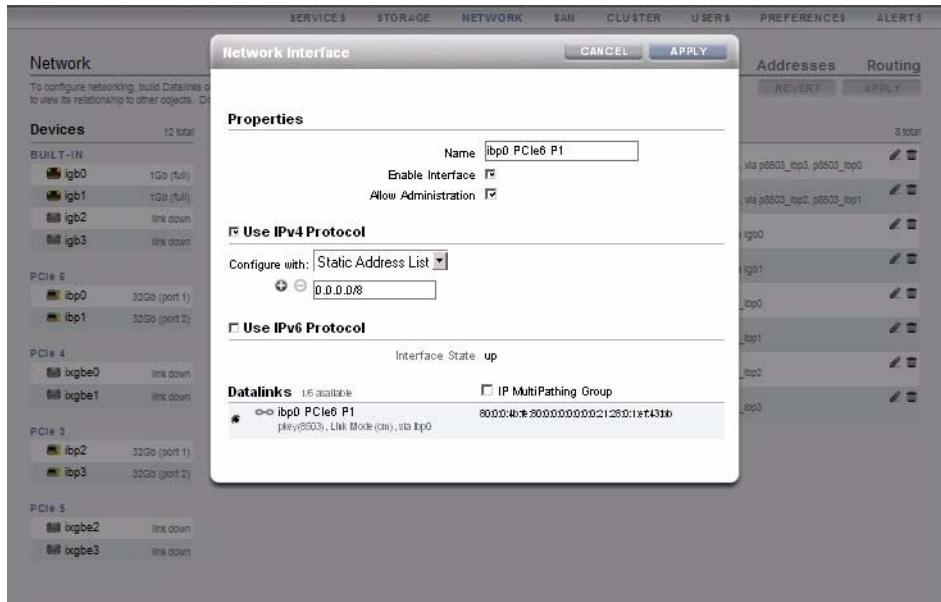


Configuring Oracle ZFS Storage ZS3-BA Networking for an Active-Active Configuration

Configure the InfiniBand ports on the Oracle ZFS Storage ZS3-BA for IP multipathing. Four IP addresses, on the private storage subnet, are needed for each Oracle ZFS Storage ZS3-BA head (therefore, eight addresses total) because the interfaces will run in an active-active configuration.

1. Configure each InfiniBand datalink as its own network interface.
 - a. Log on to the BUI of Head 1 and navigate to **Configuration > Network**.
 - b. Click the plus icon next to **Interfaces**. The Network Interface dialogue box opens.
 - c. Complete the dialogue box as follows (refer to [FIGURE 5-3](#)):
 - i. Enter a meaningful name for the network interface.
 - ii. Verify that `Enable Interface` is checked.
 - iii. Verify that `Allow Administration` is checked.
 - iv. Verify that `Use IPv4 Protocol` is checked.
 - v. Verify that the `Configure with` menu selection is **Static Address List**.
 - vi. In the box below that, enter **0.0.0.0/8**.
 - vii. Verify that `Use IPv6 Protocol` is not checked.
 - viii. Select the datalink for **ibp0** and click **Apply**.
 - d. Repeat steps b and c for the remaining datalinks (**ibp1**, **ibp2**, and **ibp3**).
 - e. Repeat steps a through d on Head 2.

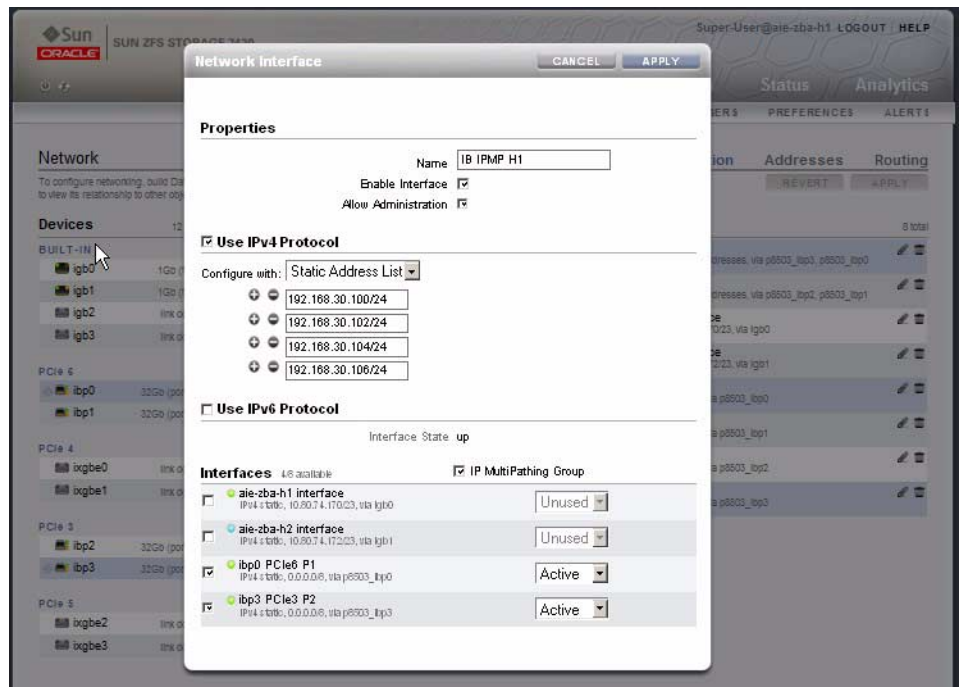
FIGURE 5-3 Creating network interfaces for each InfiniBand datalink



2. Configure the IPMP interface on Head 1.
 - a. Log on to the BUI of Head 1 and navigate to **Configuration > Network**.
 - b. Click the plus icon next to **Interfaces**. The Network Interface dialogue box opens.
 - c. Complete the dialogue box as follows (refer to [FIGURE 5-4](#)):
 - i. Enter a meaningful name for the IPMP network interface.
 - ii. Verify that **Enable Interface** is checked.
 - iii. Verify that **Allow Administration** is checked.
 - iv. Verify that **Use IPv4 Protocol** is checked.
 - v. Verify that the **Configure with** menu selection is **Static Address List**.
 - vi. Click the plus sign next to the empty box three times, so that four empty boxes are displayed.

- vii. In each empty box, enter one of the IP addresses reserved for the InfiniBand connections with its respective /24 netmask designation. As a best practice, do not use consecutive IP addresses from the block, but rather every other one (for example, all odd or all even).
- viii. Verify that Use IPv6 Protocol is not checked.
- ix. Check the IP MultiPathing Group box.
- x. Check the boxes next to the interfaces corresponding to datalinks **ibp0** and **ibp3**.
- xi. Verify that each of the two interfaces are set to **Active** and click **Apply**.

FIGURE 5-4 IPMP InfiniBand group for Oracle ZFS Storage ZS3-BA Head 1



- d. From **Configuration > Network**, click **Routing**.
- e. Click on the Multihoming model corresponding to Adaptive (refer to [FIGURE 5-5](#)).

FIGURE 5-5 Network routing for Oracle ZFS Storage ZS3-BA Head 1

SERVICES STORAGE NETWORK SAN CLUSTER USERS PREFERENCES ALERTS

Network

Configuration

Addresses

Routing

Multihoming model

Loose

Adaptive

Strict

Routing Table Entries

7 Total

ALL

:

STATIC

:

DYNAMIC

:

DHCP

:

SYSTEM

:

INACTIVE

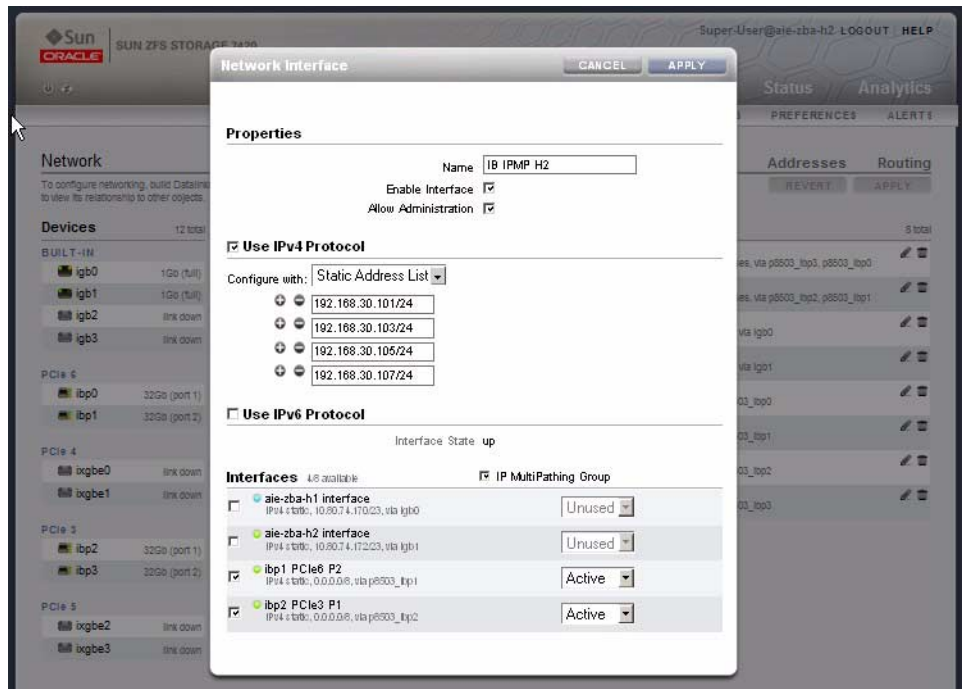
DESTINATION ▲	GATEWAY	FAMILY	TYPE	INTERFACE
default	10.80.75.254	IPv4	static	<div>aie-zba-h1 interface</div> <div>IPv4 static, 10.80.74.170/23, via igb0</div>
default	10.80.75.254	IPv4	inactive	<div>aie-zba-h2 interface</div> <div>IPv4 static, 10.80.74.172/23, via igb1</div>
10.80.74.0/23	10.80.74.170	IPv4	system	<div>aie-zba-h1 interface</div> <div>IPv4 static, 10.80.74.170/23, via igb0</div>
192.168.30.0/24	192.168.30.106	IPv4	system	<div>IB IPMP H1</div> <div>IPMP, IPv4 static, 4 addresses, via p8503_ibp3, p8503_ibp0</div>
192.168.30.0/24	192.168.30.104	IPv4	system	<div>IB IPMP H1</div> <div>IPMP, IPv4 static, 4 addresses, via p8503_ibp3, p8503_ibp0</div>
192.168.30.0/24	192.168.30.102	IPv4	system	<div>IB IPMP H1</div> <div>IPMP, IPv4 static, 4 addresses, via p8503_ibp3, p8503_ibp0</div>
192.168.30.0/24	192.168.30.100	IPv4	system	<div>IB IPMP H1</div> <div>IPMP, IPv4 static, 4 addresses, via p8503_ibp3, p8503_ibp0</div>

3. Configure the IPMP interface on Head 2.

- a. Log on to the BUI of Head 2 and navigate to **Configuration > Network**.
- b. Click the plus icon next to **Interfaces**. The Network Interface dialogue box opens.
- c. Complete the dialogue box as follows (refer to [FIGURE 5-6](#)):
 - i. Enter a meaningful name for the IPMP network interface.
 - ii. Verify that `Enable Interface` is checked.
 - iii. Verify that `Allow Administration` is checked.
 - iv. Verify that `Use IPv4 Protocol` is checked.
 - v. Verify that the `Configure with` menu selection is **Static Address List**.
 - vi. Click the plus sign next to the empty box three times, so that four empty boxes are displayed.
 - vii. In each empty box, enter one of the remaining four IP addresses reserved for the InfiniBand connections with its respective /24 netmask designation. These should be the ones **not** used on Head 1.
 - viii. Verify that `Use IPv6 Protocol` is not checked.
 - ix. Check the `IP MultiPathing Group` box.

- x. Check the boxes next to the interfaces corresponding to datalinks **ibp1** and **ibp2**.
- xi. Verify that each of the two interfaces are set to **Active** and click **Apply**.

FIGURE 5-6 IPMP InfiniBand group for Oracle ZFS Storage ZS3-BA Head 2



- d. From **Configuration > Network**, click **Routing**.
 - e. Click on the Multihoming model corresponding to Adaptive.
4. Verify connectivity with the Oracle SuperCluster nodes.

Verify that each node can ping each of the eight addresses used in the IPMP groups on the Oracle ZFS Storage ZS3-BA. Add these IP addresses to the `/etc/inet/hosts` table of each node.

Configuring the Oracle ZFS Storage ZS3-BA Storage Pool

Pool configuration assigns physical disk drive resources to logical storage pools for backup data storage. To maximize system throughput, configure two equally sized storage pools by assigning half of the physical drives in each drive tray to each storage pool as shown in [FIGURE 5-7](#).

FIGURE 5-7 Storage pool configured based on half of drives in each tray

Configure external storage

ABORT

COMMIT

Allocate and verify storage

< Step 1 of 2 >

Before configuration, you must allocate storage to the pool. SATA storage must be added in whole or half-chassis units; storage in SAS-2 enclosures may be added on a per-device basis. You may leave some storage unallocated. This step will also verify that all devices are present and minimally functional. You can configure storage with affected devices, but they will not be available to the storage pool and cannot be added later without reconfiguring all storage. It is recommended that configuring the affected chassis be deferred until any problems can be repaired.

NAME	MODEL	ALLOCATION	DATA	LOG	CACHE	
aie-7420a-h1	Sun ZFS Storage 7420		-	-	2 ▾ (954G)	
0945QCQ004	Sun Disk Shelf (SAS-2)		12 ▾ (43.7T)	-	-	
1001QCQ032	Sun Disk Shelf (SAS-2)		10 ▾ (36.4T)	2 ▾ (68G)	-	
1001QCQ024	Sun Disk Shelf (SAS-2)		12 ▾ (43.7T)	-	-	
1001QCQ02D	Sun Disk Shelf (SAS-2)		12 ▾ (43.7T)	-	-	

The Oracle ZFS Storage ZS3-BA management software presents a warning message about efficiency when two pools with the same RAID protection profile are configured. This message can be safely ignored when configuring for a high-performance Oracle RMAN backup solution.

Configuring the Oracle ZFS Storage ZS3-BA Shares

Share configuration is the process of setting up and running NFS mount points for client access. Two projects should be created for the Oracle SuperCluster configuration: one project per pool. A project is an entity that provides a higher level management interface point for a collection of shares. To optimize share management, update the default mount point for shares contained in the project to reference the database name, such as `/export/dbname`. For a performance-optimized system, create four shares for each project in each pool, for a total of eight shares (four for each head). To configure a project, perform the following:

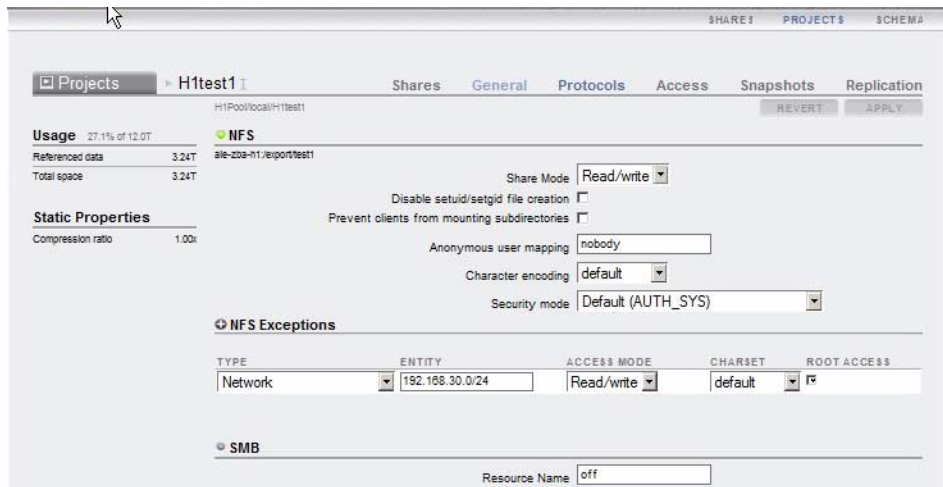
1. Log on to the BUI of Head 1 and navigate to **Shares > Projects**.
2. Click the plus icon next to **Projects**, enter a meaningful name for the project, and click **Apply**. Since a similar project will be created on the other head, uniquely name the project for Head 1, such as H1-mydb.
3. Click the pencil icon next to the new project name to edit the project.
4. Click **General** and complete the properties as follows (refer to [FIGURE 5-8](#)):
 - a. Change the Mountpoint to include the database name (for example, /export/H1-mydb).
 - b. Change Synchronous write bias from Latency to **Throughput** and click **Apply**.

FIGURE 5-8 Project general parameter settings

The screenshot displays the Oracle SuperCluster Backup User Interface (BUI) for a project named 'H1test1'. The 'General' tab is selected, showing various configuration options. The 'Usage' section indicates that 27.1% of 12.0T is used, with 3.24T of referenced data and 3.24T of total space. The 'Space Usage' section shows 'DATA' with 'Quota' and 'Reservation' both set to 0 G. The 'Static Properties' section shows a 'Compression ratio' of 1.00x. The 'Inherited Properties' section includes 'Mountpoint' set to '/export/test1', 'Read only' (unchecked), 'Update access time on read' (unchecked), 'Non-blocking mandatory locking' (unchecked), 'Data deduplication (warning)' (unchecked), 'Data compression' set to 'Off', 'Checksum' set to ' Fletcher4 (Standard) ', 'Cache device usage' set to 'All data and metadata', 'Synchronous write bias' set to 'Throughput', 'Database record size' set to '128k', 'Additional replication' set to 'Normal (Single Copy)', 'Virus scan' (unchecked), 'Prevent destruction' (unchecked), and 'Restrict ownership change' (checked). The 'Default Settings' section includes 'User' set to 'nobody', 'Group' set to 'other', 'Permissions' set to 'R W X R W X R W X', 'Volume size' set to 0 G, 'Thin provisioned' (unchecked), and 'Volume block size' set to '8k'.

5. Click **Protocols** and add an NFS exception as follows (refer to [FIGURE 5-9](#)):
 - a. Click the plus icon next to **NFS Exceptions**.
 - b. Change Type to **Network**.
 - c. Enter the subnet and netmask (for example, /24) of the InfiniBand network.
 - d. Change Access Mode to **Read/Write**.
 - e. Verify that Charset is set to **default**.
 - f. Check the Root Access box and click **Apply**.

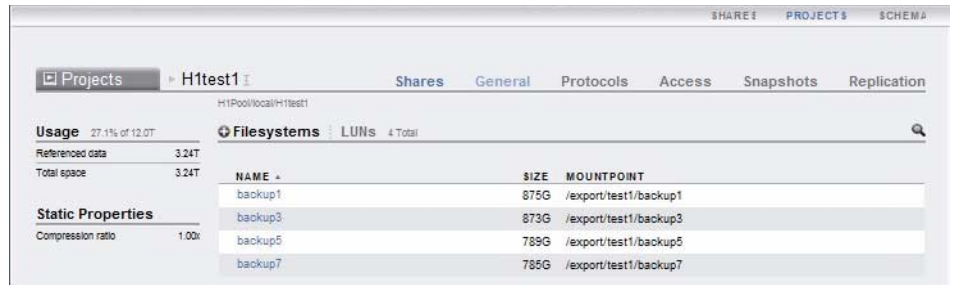
FIGURE 5-9 Setting up project NFS exceptions



6. Next to **General**, click **Shares**.
7. Create four filesystems for Head 1 and uniquely name them so they will be different from the names for Head 2. To interleave the backup streams to distribute the data across the two heads and, thereby, provide better performance, use odd-numbered names for Head 1, such as backup1, backup3, backup5, and backup7; and use even-numbered names for Head 2, such as backup2, backup4, backup6, and backup8.

To create the filesystems, click the plus icon next to **Filesystems**, enter the name of the filesystem (**backup1**), and click **Apply**. Repeat this step to create the remaining three filesystems (backup3, backup5, and backup7). The filesystem listing should be similar to [FIGURE 5-10](#).

FIGURE 5-10 Filesystem listing for Head 1



NAME	SIZE	MOUNTPOINT
backup1	875G	/export/test1/backup1
backup3	873G	/export/test1/backup3
backup5	789G	/export/test1/backup5
backup7	785G	/export/test1/backup7

- Repeat steps 1 through 7 for Head 2. Remember to use a unique project name (for example, H2-mydb), and specify even-numbered backup IDs (backup2, backup4, backup6, and backup8) for the filesystem names. The filesystem listing should be similar to FIGURE 5-11.

FIGURE 5-11 Filesystem listing for Head 2



NAME	SIZE	MOUNTPOINT
backup2	874G	/export/test1/backup2
backup4	782G	/export/test1/backup4
backup6	788G	/export/test1/backup6
backup8	754G	/export/test1/backup8

Configuring the Oracle ZFS Storage ZS3-BA DTrace Analytics

The Oracle ZFS Storage ZS3-BA includes a comprehensive performance analysis tool called DTrace Analytics. DTrace Analytics is a framework that monitors important subsystem performance accounting statistics. A subset of the available accounting statistics should be monitored to provide comprehensive data on the effectiveness and performance of Oracle RMAN backup and restore workloads.

The following Analytics are available when advanced analytics are configured on the Oracle ZFS Storage ZS3-BA (Configuration > Preferences > Enable Advanced Analytics):

- *CPU*: Percent utilization broken down by CPU mode
- *Disk*: Average number of I/O operations broken down by state of operation
- *Disk*: I/O bytes per second broken down by type of operation
- *Disk*: I/O operations per second broken down by latency
- *Disk*: Disks with utilization of at least 95 percent broken down by disk
- *Network*: Interface bytes per second broken down by direction
- *Network*: Interface bytes per second broken down by interface
- *Protocol*: NFSv3 operations per second broken down by size
- *Protocol*: NFSv3 operations per second broken down by type of operation
- *Protocol*: NFSv3 operations per second of type read broken down by latency
- *Protocol*: NFSv3 operations per second of type write broken down by latency
- *Protocol*: NFSv3 operations per second of type read broken down by size
- *Protocol*: NFSv3 operations per second of type write broken down by size

Implementing these accounting statistics helps end-users gain a quantitative understanding of the instantaneous and historical resource consumption and quality of service (QoS) for their specific implementation.

Configuring the Client NFS Mount

When configuring the Oracle ZFS Storage ZS3-BA, any server that accesses the appliance, including Oracle SuperCluster nodes, is considered a client. Configuring the client NFS mount includes creating the target directory structure for access to the Oracle ZFS Storage ZS3-BA as well as the specific NFS mount options necessary for optimal system performance. Mount options for Solaris clients are:

```
rw,bg,hard,nointr,rsz=1048576,wsz=1048576,proto=tcp,vers=3,forcedirectio
```

The mount points of the directories created on the Oracle ZFS Storage ZS3-BA should be created on each of the Oracle SuperCluster nodes and added to their `/etc/inet/hosts` table.

For detailed configuration steps, see [“Configuration Details for Client NFS Mount and Oracle dNFS”](#) on page 79.

Tuning the Solaris 11 Network and Kernel

The following entries should be added to the `/etc/system` file of each of Oracle SuperCluster node:

```
set rpcmod:clnt_max_conns = 8
set nfs:nfs3_bsize = 131072
```

Additionally, the following commands need to be run on each Oracle SuperCluster node every time it is rebooted:

```
/usr/sbin/ndd -set /dev/tcp tcp_max_buf 2097152
/usr/sbin/ndd -set /dev/tcp tcp_xmit_hiwat 1048576
/usr/sbin/ndd -set /dev/tcp tcp_recv_hiwat 1048576
```

Additional tuning might be necessary to achieve optimal performance. Refer to Oracle SuperCluster Tunables document 1474401.1, available at <http://support.oracle.com>, for the latest information. Also, the January 2013 QFSDP release added a “ssctuner” tool that automatically sets tunables. Refer to the Oracle SuperCluster release notes for additional information.

Configuring Oracle Direct NFS (dNFS)

For detailed configuration steps, see “[Configuration Details for Client NFS Mount and Oracle dNFS](#)” on page 79.

On each Oracle SuperCluster node, configure dNFS as follows:

1. Shut down the running instance of the Oracle Database software.
2. Change directory to \$ORACLE_HOME/rdbms/lib.
3. Enable dNFS:

```
make -f $ORACLE_HOME/rdbms/lib/ins_rdbms.mk dnfs_on
```
4. Update the oranfstab file (located in /\$ORACLE_HOME/dbs) with the server, path, and export names specific to the configuration, where:
 - The server parameter refers to the local name of the Oracle ZFS Storage ZS3-BA head on the InfiniBand network.
 - The path parameters should reflect the address(es) for that head specified during configuration.
 - The export parameters should reflect the mount points similar to the entries created in /etc/vfstab. The entries should look similar to the following.

For single IP configuration (only Oracle SuperCluster T5-8 without external leaf switches):

```
server: aie-zba-h1-stor
path: 192.168.30.100
export: /export/test1/backup1 mount: /zba/test1/backup1
export: /export/test1/backup3 mount: /zba/test1/backup3
```

```

export: /export/test1/backup5 mount: /zba/test1/backup5
export: /export/test1/backup7 mount: /zba/test1/backup7
server: aie-zba-h2-stor
path: 192.168.30.101
export: /export/test1/backup2 mount: /zba/test1/backup2
export: /export/test1/backup4 mount: /zba/test1/backup4
export: /export/test1/backup6 mount: /zba/test1/backup6
export: /export/test1/backup8 mount: /zba/test1/backup8

```

For IPMP Group configuration (all others):

```

server: aie-zba-h1-stor
path: 192.168.30.100
path: 192.168.30.102
path: 192.168.30.104
path: 192.168.30.106
export: /export/test1/backup1 mount: /zba/test1/backup1
export: /export/test1/backup3 mount: /zba/test1/backup3
export: /export/test1/backup5 mount: /zba/test1/backup5
export: /export/test1/backup7 mount: /zba/test1/backup7
server: aie-zba-h2-stor
path: 192.168.30.101
path: 192.168.30.103
path: 192.168.30.105
path: 192.168.30.107
export: /export/test1/backup2 mount: /zba/test1/backup2
export: /export/test1/backup4 mount: /zba/test1/backup4
export: /export/test1/backup6 mount: /zba/test1/backup6
export: /export/test1/backup8 mount: /zba/test1/backup8

```

5. Restart the Oracle Database software instance.

Tuning the Oracle Database Instance for Oracle RMAN Backup and Restore

Optimizing high-bandwidth backup and restore operations using Oracle RMAN and the Oracle ZFS Storage Appliance requires adjusting the instance parameters that control I/O buffering. For information about how to tune these parameters, see Article ID 1072545.1: *RMAN Performance Tuning Using Buffer Memory Parameters*) at <http://support.oracle.com>.

For Oracle SuperCluster, tuning the following four parameters should be considered:

- `_backup_disk_bufcnt` - Number of buffers used to process backup sets
- `_backup_disk_bufsz` - Size of the buffers used to process backup sets
- `_backup_file_bufcnt` - Number of buffers used to process image copies
- `_backup_file_bufsz` - Size of the buffers used to process image copies

For backup and restore operations on backup sets and image copies, set the number of buffers to 64 and the buffer size to 1 MB:

```
SQL> alter system set "_backup_disk_bufcnt"=64;
SQL> alter system set "_backup_file_bufcnt"=64;
SQL> alter system set "_backup_disk_bufsz"=1048576;
SQL> alter system set "_backup_file_bufsz"=1048576;
```

These commands may be configured persistently by adding them to the `SPFILE`, or they may be set dynamically in the Oracle RMAN run block used to execute the backup or restore operations.

The following code fragments show how to dynamically tune the buffer sizes and counts for backup and restore operations.

- Backup set backup:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    allocate channel...
    ...
    backup as backupset database;
}
```

- Backup set restore:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    allocate channel...
    ...
    restore database;
}
```

- Image copy backup:

```
run
{
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...
```

```
...
    backup as copy database;
}
```

■ Image copy restore:

```
run
{
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...
...
    restore database;
}
```

Performing an incrementally applied backup requires reading an incremental backup set and writing to an image copy. To tune buffers for incrementally applied backups, run the following:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64';
    sql 'alter system set "_backup_disk_bufsz"=1048576';
    sql 'alter system set "_backup_file_bufcnt"=64';
    sql 'alter system set "_backup_file_bufsz"=1048576';
    allocate channel...
...
    recover copy of database;
}
```

Creating Dedicated Services for Oracle RMAN Operations

Two services dedicated to Oracle RMAN processing can be configured to optimize management of load balancing, high availability, and upgrades. These services can be evenly load balanced over all the nodes of an Oracle SuperCluster system. Availability and performance can be optimized by configuring the services to run on a preferred instance while preparing them to fail over to any instance in the cluster. If these services are configured, upgrading a one-quarter or one-half rack Oracle SuperCluster system does not require changing the connect string of the Oracle RMAN run block.

The `srvctl` utility is used to install services for Oracle RMAN processing. The following code fragment shows how to create two services evenly distributed over a four-node cluster that are set up to fail over to any other node in the cluster. In this example, the services are installed for a database named `dbname` and are named `dbname_bkup[1-2]`.

```
srvctl add service -d dbname -r dbname1 -a dbname2 -s dbname_bkup1
srvctl start service -d dbname -s dbname_bkup1
srvctl add service -d dbname -r dbname2 -a dbname1 -s dbname_bkup2
srvctl start service -d dbname -s dbname_bkup2
```

Configuring Oracle RMAN

Configuring Oracle RMAN channel and parallelism includes specifying the file system targets for the Oracle RMAN backup channels and the total number of channels used for backup and restore operations. Performance benefits can be realized by configuring 16 Oracle RMAN channels spanning the available Oracle ZFS Storage ZS3-BA shares. Configure Oracle RMAN channels such that they are evenly distributed over the Oracle Database instances and nodes in the RAC cluster and evenly distributed over the shares exported from the Oracle ZFS Storage ZS3-BA.

The following code fragments show sample Oracle RMAN run blocks for performing backup and restore operations for backup sets and image copies as well as applying incremental merges to image copies. The sample code is based on the following database configuration:

- Database name: `dbname`
- SYSDBA login: `sys/welcome`
- Scan address: `ad01-scan`
- Service names for the backup: `dbname_bkup[1-2]`

The Oracle ZFS Storage ZS3-BA can be configured in a one-pool configuration in which the appliance exports eight shares used as eight mount points.

The Oracle RMAN run blocks for backup and restore using backup sets and image copies are shown in the examples in the sections below. In these examples, the mount points for the four-share configuration are accessed as `/zfssa/dbname/backup1` through `/zfssa/dbname/backup4`. Also, the examples are for a configuration in which the Oracle ZFS Storage ZS3-BA exports four shares used as four mount points for 16 Oracle RMAN channels.

Backup set level 0 backup:

```
run
{
    sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
    sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';

    allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
    allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
    allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';
    allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';
    allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
    allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
    allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';
    allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';

    allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
    allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
    allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';
    allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';
    allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
    allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
    allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';
    allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';

    configure snapshot controlfile name to
    '/zfssa/dbname/backup1/snapcf_dbname.f';

    backup as backupset incremental level 0 section size 32g database
    tag 'FULLBACKUPSET_L0' plus archivelog tag 'FULLBACKUPSET_L0';
}
```

Backup set level 1 backup:

```
run
{
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';

  allocate channel ch09 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';
  allocate channel ch13 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch14 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
  allocate channel ch15 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';
  allocate channel ch16 device type disk connect 'sys/welcome@ad01-
  scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';

  configure snapshot controlfile name to
  '/zfssa/dbname/backup1/snapcf_dbname.f';
```

```

backup as backupset incremental level 1 database tag
'FULLBACKUPSET_L1' plus archivelog tag 'FULLBACKUPSET_L1';
}

```

Image copy backup:

```

run
{
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';

  allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';
  allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';
  allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';
  allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';
  allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';
}

```

```

configure snapshot controlfile name to
'/zfsa/dbname/backup1/snapcf_dbname.f';

backup incremental level 1 for recover of copy with tag 'IMAGECOPY'
database;
}

```

Incremental merge to image copy:

```

run
{
    sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
    sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
    sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
    sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';

    allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
    allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
    allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
    allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';

    allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
    allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
    allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
    allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
}

```

```

allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

recover copy of database with tag 'IMAGECOPY';
}

```

Restore validate:

```

run
{
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';

  allocate channel ch01 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';

  allocate channel ch09 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
}

```

```

allocate channel ch13 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
allocate channel ch14 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';
allocate channel ch15 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup1';
allocate channel ch16 device type disk connect 'sys/welcome@ad01-
scan/dbname_bkup2';

configure snapshot controlfile name to
'/zfssa/dbname/backup1/snapcf_dbname.f';

restore validate database;
}

```


Configuration Details for Client NFS Mount and Oracle dNFS

This chapter contains sample scripts showing how to attach the Oracle ZFS Storage ZS3-BA to an Oracle Exadata or Oracle SuperCluster. These scripts are designed to support a database named `dbname` in a one-pool and a two-pool Oracle ZFS Storage ZS3-BA configuration.

General Implementation Steps

The implementation steps are:

1. Set up the directory structure (mount points) to mount the shares on the host.
2. Update a file to mount the exported shares to the appropriate mount points.
 - **Oracle Exadata:** Update `/etc/fstab` to mount the shares exported from the Oracle ZFS Storage ZS3-BA to the appropriate mount points.
 - **Oracle SuperCluster:** Update `/etc/vfstab` to mount the shares exported from the Oracle ZFS Storage ZS3-BA to the appropriate mount points.
3. Automate mounting and unmounting of the shares.
 - **Oracle Exadata:** Create an `init.d` service to automate the process of mounting and unmounting the shares.
 - **Oracle SuperCluster:** Enable the NFS client services to mount the NFS shares at reboot.
4. Update the `oranfstab` file to access the Oracle ZFS Storage ZS3-BA exported shares or set mount on boot in `/etc/fstab`.
5. Mount the shares on the host.

6. Change the permissions of the mounted shares to match the permission settings of ORACLE_HOME.
7. Restart the Oracle Database instance to pick up the changes to the oranfstab file.

These steps are described in more detail in the following sections.

Note – If you used the Oracle Engineered Systems Backup Utility, all steps except for step 4 and step 7 have already been performed for you.

Detailed Implementation Steps

Setting Up the Directory Structure to Mount the Shares on the Host

Set up mount points for the shares on the host as shown.

```
mkdir -p /zfssa/dbname/backup1
mkdir -p /zfssa/dbname/backup2
mkdir -p /zfssa/dbname/backup3
mkdir -p /zfssa/dbname/backup4
```

Updating a File to Mount Exported Shares

Follow the procedure in the appropriate subsection for your system:

- [“Updating the /etc/fstab File for Oracle Exadata” on page 81](#)
- [“Updating the /etc/vfstab File for Oracle SuperCluster” on page 82](#)

Updating the /etc/fstab File for Oracle Exadata

To update the /etc/fstab file, use the appropriate following option.

Note – The UNIX newline escape character (\) indicates a single line of code has been wrapped to a second line in the listing below. When entering a wrapped line into fstab, remove the \ character and combine the two line segments into a single line.

For a one-pool configuration:

```
192.168.36.200:/export/dbname/backup1 /zfssa/dbname/backup1 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0

192.168.36.200:/export/dbname/backup2 /zfssa/dbname/backup2 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0

192.168.36.200:/export/dbname/backup3 /zfssa/dbname/backup3 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0

192.168.36.200:/export/dbname/backup4 /zfssa/dbname/backup4 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0
```

For a two-pool configuration:

```
192.168.36.200:/export/dbname/backup1 /zfssa/dbname/backup1 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0

192.168.36.201:/export/dbname/backup2 /zfssa/dbname/backup2 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0

192.168.36.200:/export/dbname/backup3 /zfssa/dbname/backup3 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0

192.168.36.201:/export/dbname/backup4 /zfssa/dbname/backup4 nfs \
noauto,rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,\
nfsvers=3,timeo=600 0 0
```

Updating the /etc/vfstab File for Oracle SuperCluster

To update the /etc/vfstab file, use the appropriate following option.

Note – The UNIX newline escape character (\) indicates a single line of code has been wrapped to a second line in the listing below. When entering a wrapped line into vfstab, remove the \ character and combine the two line segments into a single line.

For a one-pool configuration:

```
192.168.36.200:/export/dbname/backup1 - /zfssa/dbname/backup1 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio

192.168.36.200:/export/dbname/backup2 - /zfssa/dbname/backup2 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio

192.168.36.200:/export/dbname/backup3 - /zfssa/dbname/backup3 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio

192.168.36.200:/export/dbname/backup4 - /zfssa/dbname/backup4 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio
```

For a two-pool configuration:

```
192.168.36.200:/export/dbname/backup1 - /zfssa/dbname/backup1 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio

192.168.36.201:/export/dbname/backup2 - /zfssa/dbname/backup2 \
nfs - yes \ rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio

192.168.36.200:/export/dbname/backup3 - /zfssa/dbname/backup3 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio

192.168.36.201:/export/dbname/backup4 - /zfssa/dbname/backup4 \
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto=\
tcp,vers=3,forcedirectio
```

Automating Share Mounting and Unmounting

Follow the procedure in the appropriate subsection for your system:

- [“Creating an init.d Service for Oracle Exadata” on page 83](#)
- [“Enabling the NFS Client Service for Oracle SuperCluster” on page 84](#)

Creating an init.d Service for Oracle Exadata

Create an `init.d` service using the appropriate following option.

```
#!/bin/sh
#
# zfssa_dbname: Mount ZFSSA project dbname for database dbname
#
# chkconfig: 345 61 19
# description: mounts ZFS Storage Appliance shares
#

start()
{
    mount /zfssa/dbname/backup1
    mount /zfssa/dbname/backup2
    mount /zfssa/dbname/backup3
    mount /zfssa/dbname/backup4
    echo "Starting $prog: "
}

stop()
{
    umount /zfssa/dbname/backup1
    umount /zfssa/dbname/backup2
    umount /zfssa/dbname/backup3
    umount /zfssa/dbname/backup4
    echo "Stopping $prog: "
}

case "$1" in
    start)
        start
        ;;
    stop)
        stop
        ;;
    *)
        ;;
esac
```

```

restart)
    stop
    start
    ;;
status)
    mount
    ;;
*)
    echo "Usage: $0 {start|stop|restart|status}"
    exit 1
esac

```

(Optional) Enable the `init.d` service for start-on-boot by entering:

```
# chkconfig zfssa_dbname on
```

(Optional) Start and stop the service manually using the service commands:

```
# service zfssa_dbname start
# service zfssa_dbname stop
```

Enabling the NFS Client Service for Oracle SuperCluster

Enable the NFS Client Service on the Solaris 11 host with the following command:

```
svcadm enable -r nfs/client
```

Updating oranfstab to Access Oracle ZFS Storage ZS3-BA Exports

Follow the procedure in the appropriate subsection for your system:

- [“Updating oranfstab to Access Oracle ZFS Storage ZS3-BA Exports for Oracle Exadata” on page 85](#)
- [“Updating oranfstab to Access Oracle ZFS Storage ZS3-BA Exports for Oracle SuperCluster” on page 85](#)

Updating oranfstab to Access Oracle ZFS Storage ZS3-BA Exports for Oracle Exadata

Note – If you used the Oracle Engineered Systems Backup Utility, this procedure has **not** been performed for you, and you must complete this procedure yourself.

To update the oranfstab file to access Oracle ZFS Storage ZS3-BA exports for Oracle Exadata, use the appropriate following option.

For a one-pool configuration:

```
server: 192.168.36.200
local: 192.168.36.100 path: 192.168.36.200
local: 192.168.36.101 path: 192.168.36.201
local: 192.168.36.102 path: 192.168.36.202
local: 192.168.36.103 path: 192.168.36.203
export: /export/dbname/backup1 mount: /zfssa/dbname/backup1
export: /export/dbname/backup2 mount: /zfssa/dbname/backup2
export: /export/dbname/backup3 mount: /zfssa/dbname/backup3
export: /export/dbname/backup4 mount: /zfssa/dbname/backup4
```

For a two-pool configuration:

```
server: 192.168.36.200
local: 192.168.36.100 path: 192.168.36.200
local: 192.168.36.101 path: 192.168.36.202
export: /export/dbname/backup1 mount: /zfssa/dbname-2pool/backup1
export: /export/dbname/backup3 mount: /zfssa/dbname-2pool/backup3
server: 192.168.36.201
local: 192.168.36.102 path: 192.168.36.201
local: 192.168.36.103 path: 192.168.36.203
export: /export/dbname/backup2 mount: /zfssa/dbname-2pool/backup2
export: /export/dbname/backup4 mount: /zfssa/dbname-2pool/backup4
```

Updating oranfstab to Access Oracle ZFS Storage ZS3-BA Exports for Oracle SuperCluster

Note – If you used the Oracle Engineered Systems Backup Utility, this procedure has **not** been performed for you, and you must complete this procedure yourself.

To update the oranfstab file to access Oracle ZFS Storage ZS3-BA exports for Oracle SuperCluster, use the appropriate following option.

For a one-pool configuration:

```
server: 192.168.36.200
path: 192.168.36.200
path: 192.168.36.201
path: 192.168.36.202
path: 192.168.36.203
export: /export/dbname/backup1 mount: /zfssa/dbname/backup1
export: /export/dbname/backup2 mount: /zfssa/dbname/backup2
export: /export/dbname/backup3 mount: /zfssa/dbname/backup3
export: /export/dbname/backup4 mount: /zfssa/dbname/backup4
```

For a two-pool configuration:

```
server: 192.168.36.200
path: 192.168.36.200
path: 192.168.36.202
export: /export/dbname/backup1 mount: /zfssa/dbname-2pool/backup1
export: /export/dbname/backup3 mount: /zfssa/dbname-2pool/backup3
server: 192.168.36.201
path: 192.168.36.201
path: 192.168.36.203
export: /export/dbname/backup2 mount: /zfssa/dbname-2pool/backup2
export: /export/dbname/backup4 mount: /zfssa/dbname-2pool/backup4
```

Mounting the Shares on the Host

Follow the procedure in the appropriate subsection for your system:

- [“Mounting the Shares on Oracle Exadata” on page 86](#)
- [“Mounting the Shares on Oracle SuperCluster” on page 87](#)

Mounting the Shares on Oracle Exadata

To mount the shares on the host, enter one of the following options:

```
# service mount_dbname start
```

or

```
# dcli -l root -g /home/oracle/dbs_group service mount_dbname start
```

Mounting the Shares on Oracle SuperCluster

Using the standard Solaris mount command, manually mount the shares:

```
# mount /zfssa/dbname/backup1
# mount /zfssa/dbname/backup2
# mount /zfssa/dbname/backup3
# mount /zfssa/dbname/backup4
```

Setting the Ownership of the Mounted Shares

Change the permission settings of the mounted shares to match the permission settings of ORACLE_HOME. In this example, the user and group ownerships are set to oracle:dba.

Note – If you used the Oracle Engineered Systems Backup Utility, step 1 has already been performed for you.

1. For Oracle Exadata, enter one of the following two options. For Oracle SuperCluster, enter the first option.

```
# chown oracle:dba /zfssa/dbname/*
```

or

```
# dcli -l root -g /home/oracle/dbs_group chown oracle:dba
/zfssa/dbname/*
```

2. Restart the Oracle Database instance to pick up the changes that were made to the oranfstab file using one of the following options:

- Restart one instance at a time (rolling upgrade), for example:

```
$ srvctl stop instance -d dbname -i dbname1
$ srvctl start instance -d dbname -i dbname1
$ srvctl stop instance -d dbname -i dbname2
$ srvctl start instance -d dbname -i dbname2
$ srvctl stop instance -d dbname -i dbname3
$ srvctl start instance -d dbname -i dbname3
$ srvctl stop instance -d dbname -i dbname4
$ srvctl start instance -d dbname -i dbname4
$ srvctl stop instance -d dbname -i dbname5
$ srvctl start instance -d dbname -i dbname5
$ srvctl stop instance -d dbname -i dbname6
$ srvctl start instance -d dbname -i dbname6
$ srvctl stop instance -d dbname -i dbname7
$ srvctl start instance -d dbname -i dbname7
```



```
$ srvctl stop instance -d dbname -i dbname8  
$ srvctl start instance -d dbname -i dbname8
```

- Restart the entire database, for example:
\$ srvctl stop database -d dbname
\$ srvctl start database -d dbname